Inflation Dynamics and the Role of Inflation Expectation Formation
Three Empirical Essays

Inaugural–Dissertation
zur Erlangung des Grades Doctor oeconomiae publicae (Dr. oec. publ.)
an der Ludwig-Maximilians-Universität München

2008

vorgelegt von
Steffen Henzel

Referent: Prof. Dr. Gebhard Flaig

Korreferent: Prof. Dr. Joachim Winter

Promotionsabschlussberatung 16. Juli 2008

Eidesstattliche Versicherung

Ich versichere hiermit eidesstattlich, dass ich die vorliegende Arbeit selbständig und ohne fremde Hilfe verfasst habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sowie mir gegebene Anregungen sind als solche kenntlich gemacht.
Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

München, den 25. Februar 2008

Danksagung

An dieser Stelle möchte ich die Gelegenheit nutzen, all denen meinen Dank auszusprechen, die zum Gelingen dieser Arbeit beigetragen haben. Besonderer Dank gilt meinem Betreuer *Prof. Dr. Gebhard Flaig*, der nicht nur durch seine Politik der offenen Tür ein optimales Arbeitsumfeld gewährleistet hat und dessen Hinweise mir stets neue Perspektiven eröffnet haben. Außerdem gebührt mein Dank *Prof. Dr. Joachim Winter* für konstruktive Anmerkungen während seiner Tätigkeit als Zweitgutachter.

Die ersten beiden Kapitel der vorliegenden Arbeit wären nicht möglich gewesen, ohne die gute und intensive Zusammenarbeit mit meinem Koautor *Timo Wollmershäuser*. Profitiert habe ich auch vom produktiven Gedankenaustausch in der Konjunkturabteilung des ifo Instituts. Insbesondere die Anregungen meiner Kollegen *Oliver Hülsewig*, *Johannes Mayr* und *Dirk Ulbricht* haben einen wichtigen Beitrag zum Gelingen der Doktorarbeit geleistet. Für wissenschaftlichen Ansporn danke ich meinen Doktorandenkollegen *Oliver Röhn* und *Stefan Lachenmaier*, dessen Rat auch bei technischen Fragen hilfreich war.

Nicht zuletzt bin ich meiner Familie dankbar für die uneingeschränkte Unterstützung, die ich all die Jahre erfahren habe. Meiner Lebensgefährtin *Claudia Packhäuser* danke ich für den liebevollen Rückhalt und ihre Geduld.

Contents

List of Figures

List of Tables

Preface

When making decisions today, people generally care about the future. This inevitably introduces expectations into most economic models where they exert an influence on the dynamics of the economy. However, it may also be reasonable to assume that the course of the economy influences expectations. But how do projections of future variables emerge over time? In many economic applications, it has become standard to assume rationality of private agents. The implication is that agents form expectations as the mathematical conditional expectation under complete information about all relevant variables. As a result, agents are characterized by perfect foresight within the framework of a certain model. Empirical research, however, casts serious doubts on this assumption. There may be extended periods of over- or underprediction on the part of private agents (biasedness) and, as a result, forecast errors are mostly found to be informative (inefficiency). Similarly, Estrella and Fuhrer (2002) argue that, from a theoretical point of view, models assuming rational expectations induce counterfactual dynamics. In commonly employed models, expectations would jump instantaneously after agents have observed a shock to the system and, consequently, expectations are not persistent. As a result, in purely forward–looking models, disinflation is not accompanied by output loss. This also is at odds with empirical findings (see Ball, 1995).

The three chapters of the present dissertation intend to shed some light on specific issues concerning the formation of inflation expectations. In their seminal papers, Taylor (1980), and Calvo (1983) have introduced the forward looking component of inflation dynamics. This has lead to the derivation of the forward–looking New Keynesian Phillips curve, which has then been extended to a hybrid version by Galí and Gertler (1999) also capturing a backward–looking component. Generally, expectations about the future inflation rate determine the dynamics of the inflation rate today. Most important, they influence the transmission of monetary policy actions. Hence, monetary policy aims at influencing expectations and central banks have put much emphasis on transparency, accountability and credibility. The reduction and stabilization of long–term inflation expectations – i.e. the unconditional mean of the inflation rate – is crucial for the implementation and maintenance of a low inflation regime. This is one reason why monetary policy strategies such as inflation forecast targeting have been developed, which aim at directly steering inflation expectations in order to prevent output loss (see Svensson, 1997). Therefore, it will be of importance to measure inflation expectations correctly and to analyze the expectation formation process thoroughly.

Chapter 1 presents an enhancement of the traditional techniques applied to obtain quantitative measures of inflation expectations from qualitative surveys[1]. The original method proposed by Carlson and Parkin (1975) assumes that individual survey responses on expected inflation are based on the respondents' underlying subjective probability density function. Hence, they report inflation to go up (down) if the median of this distribution lies above (below) some indifference interval. The upper and lower boundary marks the so–called just noticeable difference. Carlson and Parkin (1975) assumed symmetric and time–invariant boundaries. Additionally, they had to impose unbiasedness of expectations to obtain an estimate of the indifference interval which, of course, prevents us from testing the nature of expectation formation. Here, we can avoid this problem by using additional information from respondents of the CESifo World Economic Survey. Moreover, we can relax the assumption of a symmetric indifference interval, as it turns out that respondents seem to react more sensitively to an expected fall in the inflation rate than to a rise. Furthermore, we can reject the assumption of constant boundaries and establish a relationship from psychophysics (Weber–Fechner Law) instead: The boundaries turn out to be an increasing function of the perceived current rate of inflation. When testing the derived quantitative measures of inflation expectations, it becomes apparent that the unbiasedness assumption does not hold for all countries considered here. Thus, traditional conversion techniques can be misleading.

Chapter 2 sheds some light on the dynamics of inflation rates, explicitly allowing for non–rational expectation formation in the estimation of the New Keynesian Phillips curve (NKPC)[2]. In general, it will matter for the dynamics of the inflation rate whether expectations are rational or not. As argued above, it seems reasonable to reject rationality of expectations. Here, using the concept of subjective expectations of firms, we assume that inflation expectations, that enter the NKPC, themselves are sluggish and, hence, may introduce additional persistence to inflation rates. In empirical studies, forward–looking relationships are mostly estimated by GMM, which relies on the fact that forecast errors are orthogonal to the information set which is available to the forecaster (McCallum, 1976). In other words, GMM is only appropriate if expectations are rational in the sense of Muth (1961)

[1]The chapter relies on Henzel and Wollmershäuser (2005), which is published in *Journal of Business Cycle Measurement and Analysis.*

[2]The chapter relies on Henzel and Wollmershäuser (2008), which is forthcoming in *Economic Modelling.*

(see, Pesaran, 1987 and Mavroeidis, 2005). However, as mentioned before, inflation expectations are often found to be biased and inefficient predictors of future inflation. Hence, we estimate the NKPC by using direct measures of expectations obtained from the CESifo World economic survey. Assuming subjective instead of rational expectations, it can be shown that there is a prominent role for backward–looking behavior in the price setting of firms. Moreover, the use of survey data gives estimates that are stable over time and endogeneity of expectations can be rejected.

Chapter 3 analyzes whether the expectation formation process can be modeled by a learning rule, according to which unobserved trend inflation is estimated by signal extraction[3]. Having stated the non–rationality of inflation expectations in surveys, it becomes apparent that also the implied jump behavior of expectations following an unanticipated policy shock is at odds with the empirical observation. Hence, coming from a theoretical point of view, it makes sense to model sluggish adjustment of inflation expectations after a monetary policy shock. One reasonable assumption in this context is that trend inflation is not directly observable by private agents. A possible solution to the problem is the estimation of unobserved trend components by Kalman filter recursions. In the analysis, the latter constitute the learning rule on the part of private agents. Assuming adaptive learning behavior in the event of unobserved policy shocks will now induce real short–term effects of monetary policy, very much like repeated unanticipated policy shocks under rational expectations. This, in turn, provides a rationale why disinflation may cause significant output loss. However, this assumption has to be tested empirically and therefore, a signal extraction model is fit to survey measures of U.S. inflation expectations. These in–sample results suggest rather slow learning of trends, which would explain the sluggishness of U.S. inflation expectations during the FED's disinflation policy under the presidency of Volcker. Furthermore, in a forecasting simulation exercise, it turns out that learning by Kalman filtering approximates U.S. survey expectations closest. However, signal extraction behavior cannot explain the process of expectation formation perfectly. On the whole, the analysis shows, that a weighted average of different types of expectation formation with a prominent role for signal extraction behavior is well suited to explain survey measures of inflation expectations during the Volcker disinflation.

[3]The chapter relies on Henzel (2008), which is available as *ifo Working Paper 55.*

Chapter 1

Quantifying Inflation Expectations with the Carlson-Parkin Method – A Survey-based Determination of the Just Noticeable Difference

Abstract

This paper presents a new methodology for the determination of the just noticeable difference which is required for the quantification of qualitative survey data. Traditional conversion methods, such as the probability approach of Carlson and Parkin (1975), the regression method of Pesaran (1984) or the time-varying parameters model of Seitz (1988), require very restrictive assumptions concerning the properties of the just noticeable difference and the expectations formation process of survey respondents. Our methodology avoids these assumptions. The novelty lies in the way the boundaries, inside of which survey respondents expect the variable under consideration to remain unchanged, are determined. Instead of deriving this so-called just noticeable difference from the qualitative survey responses and from the statistical properties of the reference time-series, we directly queried them from survey respondents by a special question in the *CESifo World Economic Survey*. The new methodology is then applied to expectations about the future development of inflation which are included in the *CESifo World Economic Survey*.

1.1 Introduction

Expectations play a crucial role in macroeconomics. In consumption theory the life-cycle and permanent income approaches stress the role of expected future income. In New Keynesian Macroeconomics firms set prices as a mark-up over a weighted average of current and expected future nominal marginal costs. Central banks closely monitor the private sector's inflation expectations. Exchange rates and share prices depend on the expected future development of their fundamental determinants. Many other examples could be given.

In empirical work expectations on future macroeconomic variables can be treated in two ways. One is to set-up a theory on how private agents form their expectations. The current standard methodology for modeling expectations is to assume rationality of economic agents which goes back to the seminal paper of Muth (1961). Assuming rational expectations has the effect that empirical models can only be tested by putting up a joint hypothesis on the model and on the expectations' formation process simultaneously. The second way to introduce expectations into empirical models is through direct measures of expectations derived from surveys of households, firms and other economic agents (see Theil (1952) for an early paper). The advantage of survey data is that expectations are given exogenously in the context of a model, and that the nature of the expectations' formation process can be investigated separately.

This paper focuses on inflation expectations obtained from the *CESifo World Economic Survey (WES)*. So far, these variables have only been presented in the form of a qualitative balance statistic, indicating whether the majority of the polled economic experts expects the inflation rate to rise, to remain constant, or to decline by the end of the next six months. Qualitative surveys therefore only provide a direction of change for a given variable, rather than an exact figure. Even though this survey technique is quite common (see for example the *Consumer Survey* conducted by the European Commission)[1], balance statistics are often of limited use

[1]The reasons why survey participants are not directly asked to quantify their expectations can be divided into two categories. The first reason is of practical nature and has to do with incentives. Since the participation at the survey is voluntary, the completion of the questionnaire must be as simple as possible in order to not discourage respondents from participating. Typically, they are asked to forecast a broad set of macroeconomic variables (such as GDP growth, inflation, unemployment, interest rates, exchange rates, share prices, etc.) so that it would be relatively time-consuming to provide a precise quantitative estimate for all these variables. The second reason is of statistical nature. It is often claimed that qualitative surveys are less susceptible to measurement errors: "(...) to the extent that expectations are 'attitudes or states of mind' of the respondents and are not merely forecasts, methods based on the measurement of ordinal responses seem less likely to be subject to measurement errors than direct attempts at cardinal measurement of expectations" (Pesaran, 1984, p. 34).

for econometric analyses. For this reason, expectations which are collected as qualitative survey data are often converted into quantitative estimates of the variables under consideration.

The most widely used conversion method goes back to a paper by Carlson and Parkin (1975). Their method assumes that individual responses about the future value of a variable are based on the respondents' subjective probability density function. Respondents report a variable to go up or down if the median of their subjective probability distribution lies above or below an indifference interval. The upper and lower boundary of the indifference interval which mark the so-called just noticeable difference are derived from the respondents' aggregate answers and the time-series properties of past realizations of the macroeconomic variable under consideration. Most crucially, Carlson and Parkin (1975) assumed the aggregate distribution to be normal with symmetric and time-invariant boundaries that are allowed to vary across countries. Additionally, they imposed that the average value of past realizations and the average value of expectations must be equal, which is typically referred to as the unbiasedness of expectations.

As these assumptions are rather restrictive a number of authors suggested extensions and alternatives to the Carlson-Parkin method (see Nardo (2003) and the papers cited there). An important alternative was the regression approach which was introduced by Pesaran (1984). The basic idea is to use the relationship between realizations (measured by official statistics) and respondents' perceptions of the past (which is additionally queried in many surveys) and to estimate the just noticeable difference on the basis of this observable data. In order to quantify the respondents' expectations about the future development of the variable under consideration, Pesaran (1984) then used these estimates and imposed them on the qualitative expectations data. Thus, in contrast to Carlson and Parkin's probability approach, quantitative expectations calculated by the regression method are a function of a specific regression model, rather than a function of a specific probability distribution. While the regression approach accounted for the possibility of an asymmetric just noticeable difference, Seitz (1988) developed an important extension that explicitly allows for asymmetric and time-varying boundaries. Though theoretically appealing, this so-called time-varying parameters method was criticized mainly because of the way the boundaries were modeled using the Kalman filter technique. Moreover, as stressed by Batchelor and Orr (1988), both alternatives to the Carlson-Parkin method assume unbiased expectations, because the just noticeable difference is inferred from a regression of actual inflation on the respondents' perceived inflation.

The novelty of the present paper is that we convert qualitative survey responses into quantitative measures for inflation expectations without having to rely on assumptions concerning the evolution of the boundaries and the expectations formation process. In contrast to the three traditional methods (Carlson-Parkin, regression, time-varying parameters) we do not implicitly derive the just noticeable difference from the qualitative survey responses and from the statistical properties of the reference time-series, but from a special question in the July 2004 *CESifo WES* in which we directly query the respondents' boundaries of the indifference interval for a given current inflation rate. This allows us to explicitly test whether the boundaries are indeed symmetric and time-invariant, whether the just noticeable difference varies across countries as suggested by the traditional methods, and whether the inflation expectations that are computed on the basis of the queried just noticeable difference are indeed unbiased.

The remainder of the paper proceeds as follows. In section 1.2 we shortly present the *CESifo WES*. Section 1.3 gives a short summary of the traditional conversion methods, and applies them to selected countries included in the *CESifo WES*. Our proposal of a survey based determination of the just noticeable difference is presented in section 1.4. The paper concludes with a summary of the main findings.

1.2 The CESifo World Economic Survey

The *CESifo WES* assesses trends in the world economy by polling transnational as well as national organizations worldwide about economic developments in the respective country. It is conducted in co–operation of *Ifo Institute for Economic Research* in Munich and the *International Chamber of Commerce* in Paris.

The questionnaire of the *CESifo WES* is distributed four times a year (January, April, July and October). The participants are asked to give their assessment of the general economic situation and expectations regarding important macroeconomic indicators of the country they inhabit. Currently, the *CESifo WES* asks about 1100 experts in 90 countries. The survey was first conducted in 1983. A question on the expected inflation rate, which is in the focus of the present paper, was only included since July 1991. Survey participants are asked to give their expectations on the inflation rate by the end of the next six months. They indicate *UP* for an expected rise in the inflation rate, *SAME* for no change in the inflation rate and *DOWN* for an expected fall in the inflation rate.

The questionnaire therefore reveals qualitative information on the participants' expectations of the future inflation rate. The individual replies are combined for each country without weighting. The 'grading' procedure consists in giving a grade of 9 to positive replies (*UP*), a grade of 5 to indifferent replies (*SAME*) and a grade of 1 to negative replies (*DOWN*). The country average which may range from 1 to 9 is published as a balance statistic.[2] Average grades within the range of 5 to 9 indicate that a majority expects inflation to rise, whereas grades within the range of 1 to 5 reveal predominantly expectations of decreasing inflation rates. What is lacking is a precise quantitative estimate of the inflation rate that is expected on average.

1.3 Traditional Conversion Techniques

This section only gives a very short summary of the conversion techniques that are typically used in the literature. The survey articles by Nardo (2003) and Pesaran and Weale (2006) provide a more detailed overview of the issues surrounding the quantification of qualitative survey responses.

1.3.1 The Probability Approach of Carlson and Parkin

Conception

The probability approach was first employed by Theil (1952) and was rediscovered by Carlson and Parkin (1975) who used the method to construct quantitative measures for inflation expectations. It basically requires two types of ingredients: the basis of the variable under consideration and the qualitative answers of each respondent. The basis is simply the last value that is observable for the individual being asked. As the *CESifo WES* asks for the expected change in the inflation rate, the basis is the inflation rate π_t which is published for the current quarter.[3]

The qualitative answer of respondent i is a result of an individual probability distribution over the possible future values of the variable in question. The respondents are supposed to report the mean of the distribution. The individual answer is $DO_{i,t}$, if the mean of the expected value of the change in inflation by the end of time $t+k$, $E_t\Delta\pi_{i,t+k}$, is smaller than some value $a_{i,t}$ ($E_t\Delta\pi_{i,t+k} < a_{i,t}$). $E_t\Delta\pi_{i,t+k}$ is defined as $E_t\pi_{i,t+k} - \pi_t$ and is measured in percentage points. Likewise, the individual answer is $UP_{i,t}$, if $E_t\Delta\pi_{i,t+k}$ is larger than some value $b_{i,t}$ ($E_t\Delta\pi_{i,t+k} > b_{i,t}$). Finally, the individual answer is $SAME_{i,t}$, if $E_t\Delta\pi_{i,t+k}$ is within the lower and upper boundary of the indifference interval $a_{i,t}$ and $b_{i,t}$ ($a_{i,t} \leq E_t\Delta\pi_{i,t+k} \leq b_{i,t}$). Assume that the distributions are independent across respondents and that they have a common form with finite mean and variance. Further assume that the upper and lower boundaries are identical for all respondents in the population ($a_{i,t} = a_t$, $b_{i,t} = b_t$).

[2] $Balance\ Statistic = \frac{[(9\times\sum UP)+(5\times\sum SAME)+(1\times\sum DOWN)]}{(\sum UP+\sum SAME+\sum DOWN)}$

[3] In section 1.4 we will show that a publication or an information lag can be ruled out in our case.

Then the survey results as a whole can be interpreted as a sampling from some aggregate distribution.

From this follows that the percentage of the responses expecting a rise and a fall which we denote by UP_t and DO_t, respectively, converge to the associated population values:

$$\begin{aligned} 1 - UP_t &= \Phi\left(\frac{b_t - E_t\Delta\pi_{t+k}}{\sigma_{t+k}}\right) \\ DO_t &= \Phi\left(\frac{a_t - E_t\Delta\pi_{t+k}}{\sigma_{t+k}}\right) \end{aligned}$$

where Φ is the cumulative distribution function of an assumed standard normal variate, and $E_t\Delta\pi_{t+k}$ and σ_{t+k} are the mean and the standard deviation of the aggregate distribution of inflation expectations. The quantiles can be calculated as:

$$r_t = \Phi^{-1}(1 - UP_t) = \frac{b_t - E_t\Delta\pi_{t+k}}{\sigma_{t+k}}, \tag{1.1}$$

$$f_t = \Phi^{-1}(DO_t) = \frac{a_t - E_t\Delta\pi_{t+k}}{\sigma_{t+k}}. \tag{1.2}$$

After eliminating σ_{t+k} and by solving for $E_t\Delta\pi_{t+k}$ one finally obtains the following expression for inflation expectations:

$$E_t\Delta\pi_{t+k} = \frac{b_t\, f_t - \ a_t\, r_t}{f_t \ - \ r_t}. \tag{1.3}$$

A crucial step of the quantification procedure is the determination of the just noticeable difference, i.e. the upper and lower boundary of the indifference interval. Carlson and Parkin (1975) assumed symmetric and time-invariant boundaries: $c = -a_t = b_t$ for the just noticeable difference in average price. If we follow that procedure for the perceived change in inflation and assume on average that expectations are correct:

$$\frac{1}{T}\sum_{t=1}^{T} E_t\Delta\pi_{t+k} = \frac{1}{T}\sum_{t=1}^{T}(\pi_t - \pi_{t-k}). \tag{1.4}$$

Using equation (1.3) for calculating the expected change in the inflation rate and setting $a_t = -c$ and $b_t = c$, equation (1.4) becomes:

$$\sum_{t=1}^{T}\frac{c\ (f_t + r_t)}{f_t - r_t} = \ \sum_{t=1}^{T}(\pi_t - \pi_{t-k}),$$

which gives the following estimate for c :

$$\hat{c} = \left(\sum_{t=1}^{T} (\pi_t - \pi_{t-k}) \right) / \left(\sum_{t=1}^{T} \frac{f_t + r_t}{f_t - r_t} \right). \tag{1.5}$$

Application to the *CESifo WES*

The results of the quantification of qualitative inflation expectations from the *CESifo WES* using the Carlson-Parkin method are shown in figures 1.1 and 1.2. Even though the *CESifo WES* covers 90 countries, for the analyses in this paper, we only consider countries where the average number of respondents is reasonably large in order to comply with the assumption of normally distributed answers: France, Germany, Italy, Japan, the UK and the US.[4] Inflation expectations for the Euro Zone were computed using a weighted sum of responses for the individual member countries (which are all included in the *CESifo WES*) according to

$$DO_t^{EUR} = \sum_{j=1}^{J} \omega_t^j \, DO_t^j \quad \text{and} \quad UP_t^{EUR} = \sum_{j=1}^{J} \omega_t^j \, UP_t^j$$

where the index j refers to each of the J Euro zone member countries, ω_t^j are the country weights used by *Eurostat* to calculate the Harmonized Index of Consumer Prices (HICP) for the Euro Zone, and DO_t^j and UP_t^j are the fractions of respondents who indicated *DOWN* and *UP* in country j.[5] The sample period runs from 1991:2 to 2004:2 at a quarterly frequency. The charts show the expected inflation rate at t for $t+2$ ($E_t\pi_{t+2}$) together with the prevailing inflation rate at time t (π_t), which is taken from the OECD database. The balance statistic is depicted as a bar chart in the lower panel. The estimates of the just noticeable difference are shown in table 1.1.

For Germany, the estimated just noticeable difference is $\hat{c} = 0.27$. This means that an expected increase (fall) in the inflation rate of 0.27 percentage points is necessary to make the respondent indicate *UP* (*DOWN*) in the questionnaire. For the Euro Zone and Japan the estimates are somewhat higher.

In the case of France, the UK and the US $\hat{c}$ becomes negative, which implies that in figures 1.1 and 1.2 the converted values for inflation expectations turn out

[4]Countries in which the average number of answers is small are mainly smaller economies. This is due to the fact that the *CESifo WES* only asks domestic experts and the number of respondents is positively correlated with the size of the economy.

[5]Before 2002:1 there were no respondents from Luxembourg so that the weights had to be adjusted accordingly.

	$\hat{c}$		$\hat{c}$
France:	2.04	Euro Zone:	0.51
France: $\diamond$	-0.49	Japan:	0.72
Germany:	0.27	UK:	-0.46
Italy:	-8.43	UK: $\diamond$	-0.91
Italy: $\diamond$	0.65	US:	-0.23

Note: For the countries in which outliers occurred due to the conversion of inflation expectations from qualitative into quantitative data we calculated the indifference band with and without the outlier. A $\diamond$ marks the cases where the outlier was included.

Table 1.1: Estimates of c using the Carlson-Parkin method

to be in opposition to the direction of change indicated by the balance statistic. To explain this result, it is important to see that in those countries expectations have been, on average, far from being correct for many years. Take the US as an example. Throughout the first part of the sample, the balance statistic shows an expected rise in the inflation rate which is clearly in contrast to the disinflation episode at the beginning of the nineties. A more detailed analysis of the balance statistic reveals that for the US only 55% (57%) of expected rises (falls) have effectively been followed by an increase (decrease) in inflation six months later. By contrast, expectations in Germany have been more correct on average, as this share amounts to 73% for both, expected rises and falls. Thus, for the US the time series is forced into the 'correct' direction, because the Carlson-Parkin method assumes the unbiasedness of expectations. As $\hat{c}$ is used to scale the time series of expectations, it acts as a degree of freedom and turns the survey results upside down. For the UK and France, this is not so obvious.

In these two countries, the calculations suffer from another problem, which is the occurrence of outliers. This problem also appears for Italy, where we calculated a value of $\hat{c} = 0.65$. An explanation for the occurrence of outliers can be given by taking a look at the Italian microdata of the survey conducted in July 1996. Table 1.2 shows that none of the 16 respondents indicated UP or $SAME$. As the responses of the basic population are assumed to be normally distributed, no response in the UP-$SAME$-$DOWN$ categories may be the result of an insufficient sample size. Even after having corrected the data as proposed in appendix 1.A, the inflation expectation in the second quarter of 1996 still remains an outlier. And it is important to understand that this outlier has a decisive impact on the just noticeable difference when the Carlson-Parkin method is applied. A calculation of the indifference interval where this observation is dropped yields a value of $\hat{c} =$

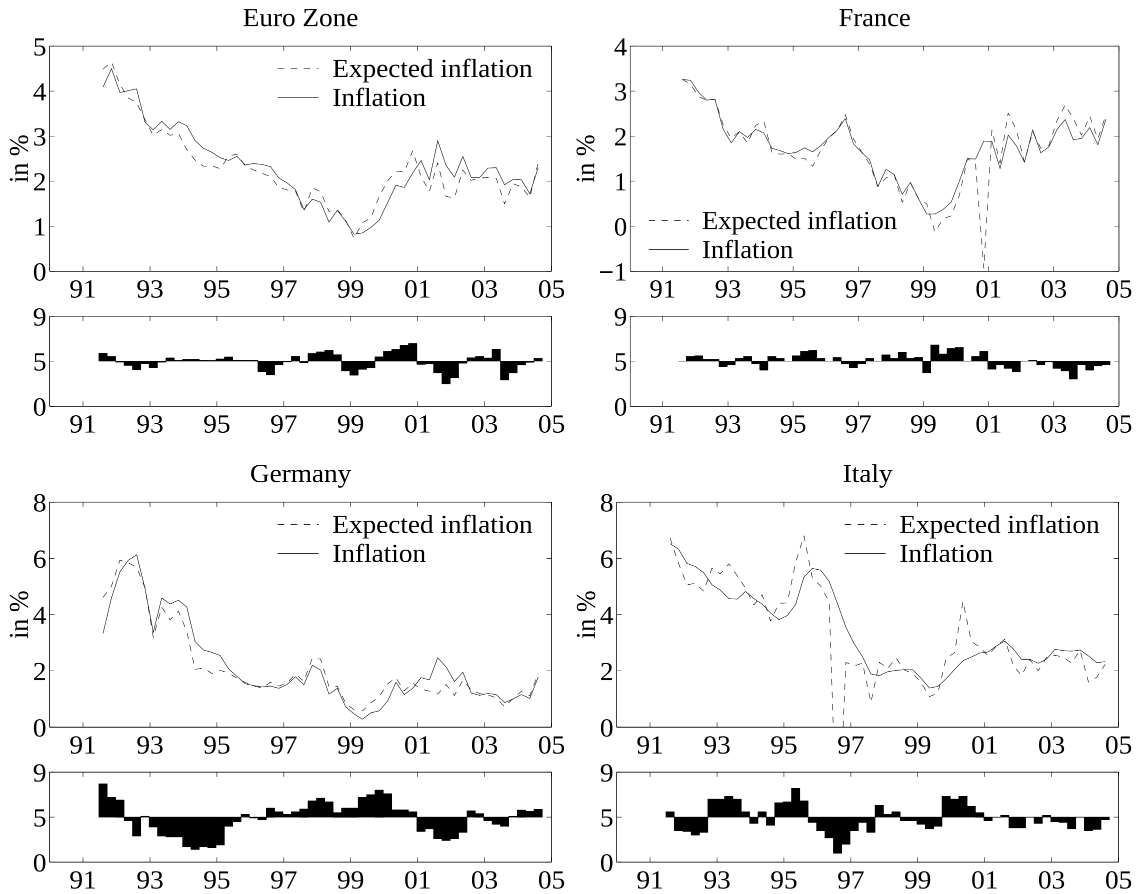


Figure 1.1: Quantification with the Carlson-Parkin method

−8.43.[6] A similar argumentation can be applied to France, where we observe an outlier in the third quarter of 2000. The outcome of the survey of October 2000 in France is shown in table 1.2. In contrast to the case before, here, the insufficient sample size results in only one respondent in the category *SAME*, while the rest is distributed over the remaining categories. Moreover, none of the corrections proposed in appendix 1.A had to be applied here. If the observation is dropped for the calculation of the indifference interval, a value for $\hat{c}$ of 2.04 results. The outcome of the survey in the UK in July 1991, where we also observe an outlier, is depicted in table 1.2. The calculation of the just noticeable difference without the outlier gives a value of −0.46. It becomes clear that the shift of the indifference interval can be quite substantial when the outlier is omitted.

[6]In table 1.1 a $\diamond$ marks the cases where the outlier was included.

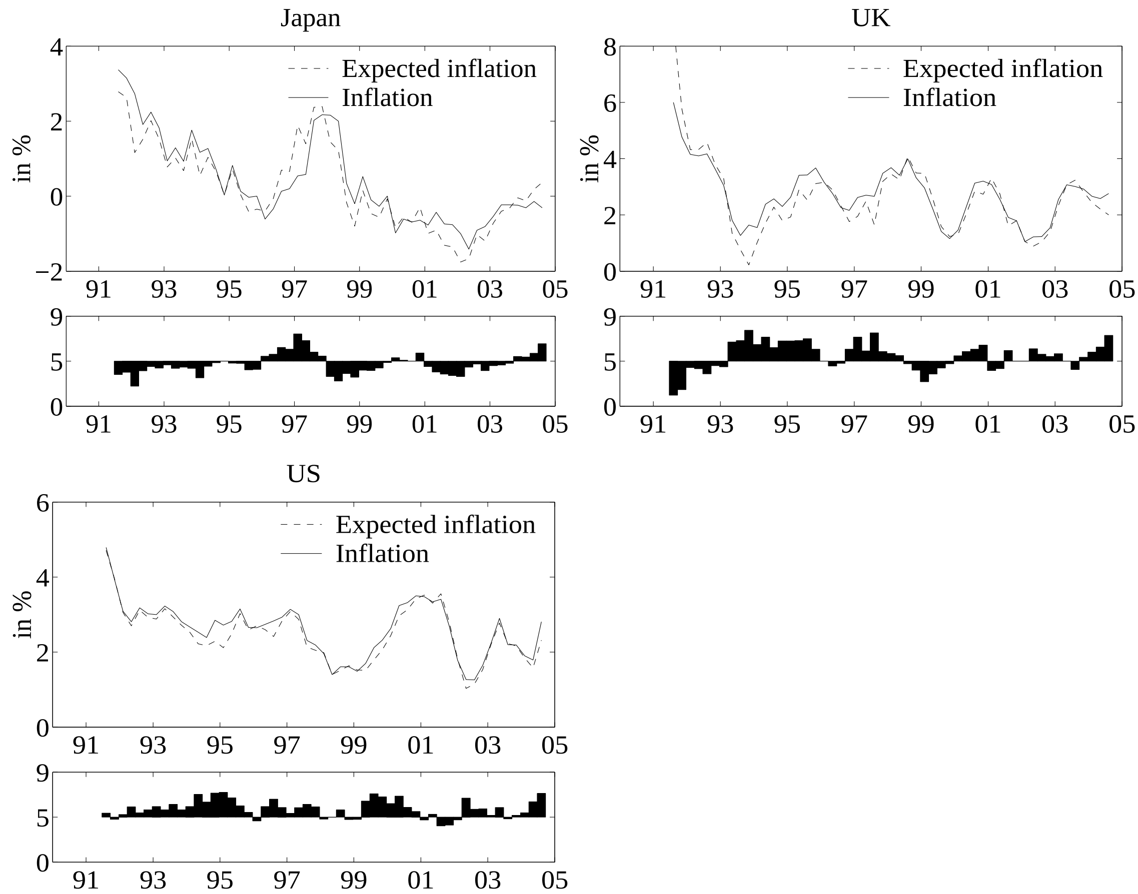


Figure 1.2: Quantification with the Carlson-Parkin method ctd.

Shortcomings of the Carlson-Parkin Method

There are several shortcomings related with the Carlson-Parkin method which have all been discussed intensively by Nardo (2003). In our view, the most important restriction is that it imposes a priori the assumption of unbiasedness, which is a necessary condition for rational expectations. Obviously, this assumption is not very useful when one wants to test the nature of the expectation formation process. For instance, 'bad expectations' are forced to be correct on average by way of scaling the time-series with the help of the indifference interval, which, in turn, can give non-interpretable results.

While Carlson and Parkin (1975) and most other subsequent papers dealing with quantification of inflation expectations resorted to scaling the quantitative series by the average level of inflation, the specific design of the inflation question in the *CESifo WES* makes it necessary to use changes in inflation instead. However,

		UP_t	$SAME_t$	DO_t
Italy (July 1996)	fractions before correction	0	0	1
	uncorrected number of responses	0	0	16
	fractions after correction	0.015	0.015	0.967
	corrected number of responses	0.24	0.24	15.52
France (Oct. 2000)	fractions before correction	0.619	0.048	0.333
	uncorrected number of responses	13	1	7
	fractions after correction	–	–	–
	corrected number of responses	–	–	–
UK (July 1991)	fractions before correction	0	0.056	0.944
	uncorrected number of responses	0	1	17
	fractions after correction	0.027	0.056	0.917
	corrected number of responses	0.49	1	16.51

Table 1.2: Violation of the normality assumption

this can create a problem if inflation has no trend, resulting in an average change of zero and, hence, $\hat{c} = 0$, irrespective of survey results f_t and r_t. Note that in the applications of the previous section, all average changes in the inflation rate $(\pi_t - \pi_{t-2})$ were below zero, ranging from -0.05 in France to -0.25 in the UK.

A further shortcoming is that the indifference interval is endogenously determined and, hence, changes with the observed survey results. It also moves with the corrections we had to make due to the violation of the normality assumption. Altogether, the fact that we calculated so many different indifference intervals across countries does not seem to be very plausible. In fact, there is no obvious reason why the perception of changes in the inflation rate should differ so dramatically across countries. Of course, given the very small average changes in inflation rates over the various samples, it is not clear, whether the estimates of c are significantly different in a statistical sense. A plausible stochastic process for generating the data could give rise to a very high variance of these estimates. One possibility to circumvent this problem would be to try to develop a method that first converts expected changes in inflation rates to levels of expected inflation before trying to scale for a just noticeable difference. We propose another possibility in section 1.4 and obtain critical values from the survey itself.

While these problems of the Carlson-Parkin method are rarely addressed in the literature[7], the two assumptions that c is constant in the sense that it neither

[7]The form of the distribution is the subject of research papers by Berk (2000) who studied symmetric and asymmetric t-distributions or Batchelor and Orr (1988) and Fishe and Lahiri (1981) who assume a logistic distribution. However, these authors did not propose any solution to small

varies over time nor with the inflationary environment, and that it is symmetric – meaning that respondents are equally sensitive to an expected rise and an expected fall of the inflation rate – have been subject to various modifications.

The Weber-Fechner Law

In this section we propose an extension of the Carlson-Parkin method that relaxes the assumption of constant and time-invariant boundaries of the indifference interval. In signal detection theory – a discipline of psychophysics – it is a well known concept that the just noticeable difference varies in proportion to the base stimulus an individual is exposed to. In other words, the higher the level of the base stimulus, the higher must be the change of this stimulus to be perceived by an individual. As this was first discovered by Weber (1834) and Fechner (1889), this concept is called the Weber-Fechner law. It was originally proven in experiments for physical stimuli like sound and weight and it has already been addressed in studies by Batchelor (1986), Batchelor and Orr (1988) and Fishe and Lahiri (1981).[8]

In this section we integrate the Weber-Fechner law into the Carlson-Parkin method. Therefore, equation (1.5) has to be modified in order to allow for variable, but still symmetric, boundaries of the indifference interval. According to the Weber-Fechner law, the just noticeable difference c can be written as a linear function of the base stimulus, which is in our case π_t:

$$c = \gamma\, \pi_t. \tag{1.6}$$

Thus, c varies over time in proportion to the inflation rate that prevails at the time expectations are formed. γ is the scaling factor which has to be computed in order to convert qualitative expectations into quantitative measures. Inserting equation (1.6) into equation (1.5) gives the following estimate for γ:

$$\hat{\gamma} = \left(\sum_{t=1}^{T} (\pi_t - \pi_{t-k}) \right) \Big/ \left(\sum_{t=1}^{T} \frac{\pi_t\,(f_t + r_t)}{f_t - r_t} \right). \tag{1.7}$$

sample problems.

[8]Batchelor (1986) calculates symmetric indifference bands with the help of the Carlson-Parkin method using qualitative survey data of eight European Community countries over the period 1974-1982. The theoretical model that he uses to describe the Weber-Fechner law is derived from the optimizing behavior of agents that minimize a statistical error. He finds that the perception of the inflation rate cannot be described by the Weber-Fechner law in its original version. Instead he estimates a negative influence of the base stimulus on the magnitude of the just noticeable difference. Nevertheless, he comes to the conclusion that the assumption of a constant indifference interval is untenable.

For the countries in our sample the calculated values for $\hat{\gamma}$ are shown in table 1.3. Similar to the results obtained by the Carlson-Parkin method, for some countries in which $\hat{\gamma} < 0$ the upper and lower boundary of the indifference interval are turned upside down. In addition to the problem of the correct sign which arises from the imposition of unbiased expectations, the parameter $\hat{\gamma}$ varies remarkably across countries for which we do not find a plausible explanation. As before, the results seem to be driven to a large extent by some single observations, because the results change significantly when the outliers are dropped. Finally, the main shortcomings of the Carlson-Parkin method are not resolved.

	$\hat{\gamma}$		$\hat{\gamma}$
France:	0.43	Euro Zone:	0.15
France:$^\diamond$	-0.65	Japan:	0.52
Germany:	0.05	UK:	-0.23
Italy:	-1.82	UK:$^\diamond$	-3.44
Italy:$^\diamond$	0.15	US:	-0.09

Note: For the countries in which outliers occurred due to the conversion of inflation expectations from qualitative into quantitative data we calculated the indifference band with and without the outlier. A $^\diamond$ marks the cases where the outlier was included.

Table 1.3: Estimates of γ

1.3.2 The Regression Approach

The OLS Method

As an alternative to the probability method for the derivation of quantitative expectations, Pesaran (1984) developed the regression approach. Rather than being a function of a specific probability distribution, the just noticeable difference and, hence, quantitative expectations are a function of a specific regression model. In contrast to the probability method, the regression approach allows for an asymmetric indifference interval. Using equation (1.3) and assuming that a and b are constant over time, the boundaries can be estimated with OLS:

$$E_t \Delta\pi_{t+k} = b\frac{f_t}{f_t - r_t} - a\frac{r_t}{f_t - r_t} + \varepsilon_t. \tag{1.8}$$

The problem however is that equation (1.8) cannot be estimated directly, since $E_t\Delta\pi_{t+k} = E_t\pi_{t+k} - \pi_t$ is unknown. In fact, it will be the outcome of the quantification procedure. One possibility to solve this problem is proposed by Entorf (1990), who replaces expectations with future realizations:[9]

$$E_t\Delta\pi_{t+k} = \Delta\pi_{t+k}. \tag{1.9}$$

Using equation (1.9) together with equation (1.8) we can estimate the unknown parameters. The results including p-values in brackets are shown in table 1.4. Newey-West adjusted standard errors were used to calculate the test statistic.

	OLS			TVP	
	$\hat{a}$	$\hat{b}$	$-\hat{a}=\hat{b}$	$\frac{1}{T}\sum_{t=1}^{T}\hat{a}_t$	$\frac{1}{T}\sum_{t=1}^{T}\hat{b}_t$
Euro Zone:	−0.28	0.19	0.49	−0.41	0.37
	[0.01]	[0.22]		(0.20)	(0.26)
France:	−0.24	0.21	0.82	−0.34	0.30
	[0.13]	[0.34]		(0.41)	(0.23)
France:$^\diamond$	0.01	−0.07	0.69	−0.29	0.25
	[0.92]	[0.45]		(0.44)	(0.22)
Germany:	−0.29	0.25	0.85	−0.27	0.21
	[0.02]	[0.09]		(0.41)	(0.32)
Italy:	−0.44	0.16	0.12	−0.35	0.00
	[0.00]	[0.15]		(0.40)	(0.29)
Italy:$^\diamond$	−0.26	0.01	0.21	−0.35	−0.01
	[0.00]	[0.92]		(0.40)	(0.30)
Japan:	−0.65	0.68	0.82	−0.63	0.54
	[0.00]	[0.00]		(0.35)	(0.17)
UK:	−0.70	0.24	0.15	−0.77	0.38
	[0.00]	[0.09]		(0.40)	(0.51)
UK:$^\diamond$	−0.55	0.19	0.22	−0.76	0.35
	[0.00]	[0.19]		(0.40)	(0.51)
US:	−0.67	0.13	0.08	−0.75	0.25
	[0.03]	[0.29]		(0.40)	(0.32)

Note: For the countries in which outliers occurred due to the conversion of inflation expectations from qualitative into quantitative data we estimated the boundaries with and without the outlier. A $\diamond$ marks the cases where the outlier was included. P-values are presented in brackets, standard errors are in parentheses. T is the number of observations.

Table 1.4: Estimates of a and b

The estimates of the upper and lower boundary have the correct signs whenever we controlled for the outlier using a dummy variable. If the outlier is included the

[9]Pesaran (1984) originally used queried data on the perceived changes of the past which he regressed on the realized inflation rate to estimate the boundaries. As the *CESifo WES* does not query the perception of the past, this approach cannot be applied.

estimated values change to some extend but the signs remain correct, except for France where both estimates have the wrong sign. But also note that they are not significant. On the whole, only for Japan significance levels are satisfactory.

Within this framework, we can actually test whether the indifference interval is symmetric by conducting a heteroscedasticity and autocorrelation consistent Wald-test on $H_0 : -a = b$. The corresponding p-values, which are shown in table 1.4, imply that only in the US asymmetric behavior seems to play a role. In all other cases we cannot reject H_0, which is probably due to the fact that – except for Germany, Japan and the UK – the estimated boundaries are insignificant. Although the boundaries still vary across countries, the differences are somewhat smaller than in the last sections. Concerning the unbiasedness issue it can easily be shown that the expectation error is identical to the estimated error term $\hat{\varepsilon}_t$.[10] Hence, the converted inflation expectations will be unbiased by assumption.

The Time-Varying Parameters Method

An important extension to the regression approach was introduced by Seitz (1988), who proposed to estimate equation (1.3) with time-varying parameters. In contrast to the regression approach in the last section, the boundaries of the indifference interval a_t and b_t are allowed to vary over time. In order to estimate the related state-space model via Kalman filter, an assumption about the stochastic process underlying the evolution of the boundaries over time has to be made. In his paper Seitz (1988) modeled them as a random walk.[11]

Like in the last section we replace the expectation term on the LHS of equation (1.3) by its realization. To calculate the just noticeable difference over time, we used the permanent component, which we obtained from the smoothed parameter estimates. Table 1.4 presents the mean of the estimates of the upper and lower boundary. Both appear to have the correct sign on average in all countries, except for Italy when we include the outlier. However, the standard deviation of the coefficients which is shown in parentheses is quite large in all cases. We take this as a hint that the indifference interval is far from being constant over time (see appendix 1.B for a chart).

A major disadvantage of this approach is, again, that the threshold values de-

[10] Note that $E_t\Delta\pi_{t+k} = \hat{b}\frac{f_t}{f_t-r_t} - \hat{a}\frac{r_t}{f_t-r_t}$ and $\Delta\pi_{t+k} - \hat{\varepsilon}_t = \hat{b}\frac{f_t}{f_t-r_t} - \hat{a}\frac{r_t}{f_t-r_t}$. Substituting for the RHS yields $\Delta\pi_{t+k} - E_t\Delta\pi_{t+k} = \hat{\varepsilon}_t$ or $\pi_{t+k} - E_t\pi_{t+k} = \hat{\varepsilon}_t$. The original proposal by Pesaran (1984) does not avoid the unbiasedness assumption either, as it extrapolates the relationship between the respondents' perception and the actual outcome to the expected evolution of the inflation rate (see also Batchelor and Orr (1988) on this point).

[11] In accordance with Pesaran (1984), Seitz (1988) originally used queried data on the perceived changes of the past.

pend on the way expectations are connected to realizations. Another criticism concerns the way the thresholds are modeled by the estimation technique. As pointed out by Nardo (2003), there are no economic or psychological reasons to suppose that individuals have an indifference interval that follows a random walk.

1.4 Survey-Based Determination of the Just Noticeable Difference

The traditional conversion methods calculate the just noticeable difference on the basis of time-series properties of realized changes in inflation (see equations (1.4) and (1.9)). To overcome the shortcomings that are related to these proceedings and their underlying assumptions, we determine the boundaries of the indifference interval by a survey. For this purpose we asked the participants of the *CESifo WES* in July 2004 an additional question with a view to get more information about the way the respondents actually form their expectations. It was put the following way:

> The following question focuses on the expectations regarding the rate of inflation (as asked in question 4 of the *WES* questionnaire).
>
> **a)** The current rate of inflation is (change of consumer prices compared to the same month previous year): ______%.
>
> **b)** The expected rate of inflation must rise above ______% to make you mark 'higher' in the *WES* questionnaire.
>
> **c)** The expected rate of inflation must fall below ______% to make you mark 'lower' in the *WES* questionnaire.

With the help of the answers to these questions, we are able to address several important issues that are related to the traditional conversion methods. First, does the Weber-Fechner law provide a valid description of the perception of changes in inflation? And if so, does this perception follow a symmetric and linear pattern? Second, is there any evidence that the just noticeable difference varies across countries? Third, what is the information set of the respondents at the time they fill in the survey and, hence, the basis on which they form their expectations? Finally, does the assumption of unbiased expectations reflect the true process of expectation formation?

1.4.1 Data Description

Before presenting the results of our analysis, we provide a short description of the responses that we received. The additional question was answered by 437 experts from 78 countries all over the world. This has the advantage of obtaining a large spectrum of perceived inflation rates. The highest inflation rate was reported with a value of 580 per cent and the lowest had a value of -1.5 per cent. The mean of the answers concerning the perceived inflation rate was 10.54 per cent with a standard deviation of about 54.39 per cent. 95 per cent of the questionnaires were returned to the *Ifo Institute* between July 05 and July 15, 2004. Unfortunately, some of the answers were missing or incomplete so that they have been of no use for the analysis:

- 10 questionnaires were incomplete because respondents did not answer part a).
- In addition, with 35 questionnaires there was no answer to both, part b) and c).
- In the sample, there is one country which we excluded because of an exceptionally high inflation rate (Zimbabwe). This reduced the sample size by 7. Our data set then included only observed inflation rates from -1.5 to 22 per cent.
- In 16 of the remaining cases, respondents either only answered part a) and c) of the question or gave an upper boundary that was below their perceived inflation rate. These answers had to be excluded when analyzing the upper boundary.
- When analyzing the lower boundary, 33 answers were of no use because either only part a) and b) have been answered or the respondent gave a lower boundary that was above the perceived inflation rate.

When all the incomplete answers and the outliers are excluded, the number of responses amounts to 352 for the estimation of the lower boundary and 369 answers were in the data set for the estimation of the upper boundary. In the following, we will denote by a the lower boundary which we obtain by subtracting for each respondent the perceived inflation rate π^p (answer given to question a)) from the answer given to question c). This procedure yields values for a that lie between -7 and 0. The upper boundary will be denoted by b and is calculated analogously as the difference between the answer given to question b) and π^p. Here, we obtain values that range between 0 and 5.

1.4.2 Estimation of the Just Noticeable Difference

According to the Weber-Fechner law expressed in equation (1.6), the just noticeable difference should be a linear function of the actual inflation rate. As we observed a large variety of inflation rates from all over the world, it is possible to estimate the relationship and to formally test whether the Weber-Fechner law holds indeed for the perception of changes in inflation.

Given the definitions of a and b, we can estimate the following two equations:

$$a = \delta_0 + \delta_1 \pi^p + \varepsilon_a \tag{1.10}$$

$$b = \gamma_0 + \gamma_1 \pi^p + \varepsilon_b, \tag{1.11}$$

where ε_a and ε_b are the errors of the regression. The results including standard errors in parentheses are summarized in table 1.5 and depicted in figures 1.3 and 1.4. A * indicates significance at the 5% level.

	$\hat{\delta}_0$	$\hat{\delta}_1$	R^2
Lower boundary a	$-0.1388^\star$ (0.0567)	$-0.1475^\star$ (0.0113)	0.3285
	$\hat{\gamma}_0$	$\hat{\gamma}_1$	R^2
Upper boundary b	$0.3288^\star$ (0.0664)	$0.1312^\star$ (0.0136)	0.2024

Table 1.5: Estimates of the just noticeable difference

As the estimated parameters are all significant, we conclude that the Weber-Fechner law holds for the perception of changes in inflation. Our results show that the upper and lower boundaries are linear functions of the inflation rate prevailing at the time the expectations are queried. Even though these boundaries are obtained from a cross-sectional estimation, we will interpret them in section 1.4.4 as evidence for a time-varying just noticeable difference with a and b depending on the inflation rate that prevails at the time expectations are formed. If, for example, perceived inflation is 3%, an expected increase of the inflation rate of 0.72 percentage points is needed to make the respondents mark *UP* in the questionnaire. By contrast, a decrease of the inflation rate of 0.58 percentage points must be expected to make the respondent mark *DOWN*. Note that Weber (1834) and Fechner (1889) originally did not allow for a constant term in their relationship of perception and base stimulus. As opposed to physical stimuli like weight and sound, there exists no situation where

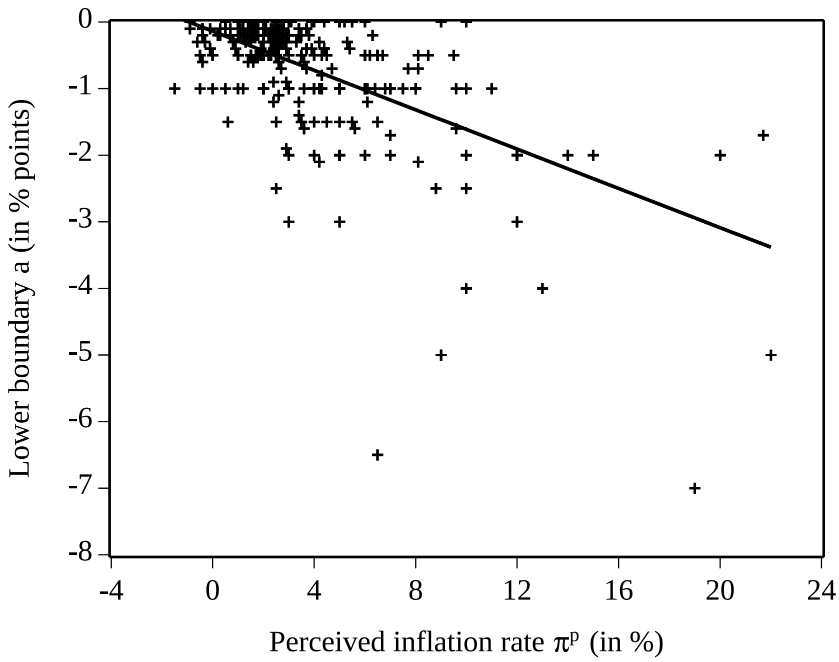


Figure 1.3: Estimation of the lower boundary

the base stimulus is not present in the case of the inflation rate. Thus, it is possible to interpret the intercept as the just noticeable difference when the perceived inflation rate is 0%. Moreover, our estimates can even be used when perceived inflation rates are negative. Specifically, the upper (lower) boundary is positive (negative) as long as π_p is greater than -2.51% (-0.94%). For values of π_p below these critical values, however, the boundaries turn upside down.

In addition to formally testing the Weber-Fechner law, the responses to our additional question can also be used to investigate whether the perception of changes in inflation indeed follows a symmetric pattern. In this context, in the literature, the assumption of a normal distribution is sometimes replaced by other asymmetric distributional assumptions like the non-central t-distribution in Berk (1999). A first hint that inflationary changes are perceived asymmetrically, is given by the fact that only about 60% of the respondents gave a symmetric indifference interval, whereas 26% (14%) gave an upper value that was larger (smaller) than the lower boundary in absolute values. The fact that there were more respondents indicating a larger upper value is reflected in our finding that $|\gamma_0| > |\delta_0|$. For the boundaries of the indifference interval, this finding together with the non-zero constant implies that $b > -a$ as long as perceived inflation is lower than 11.66%. However, we can't conclude from this, that people react less sensitively to an expected rise in the inflation rate than to

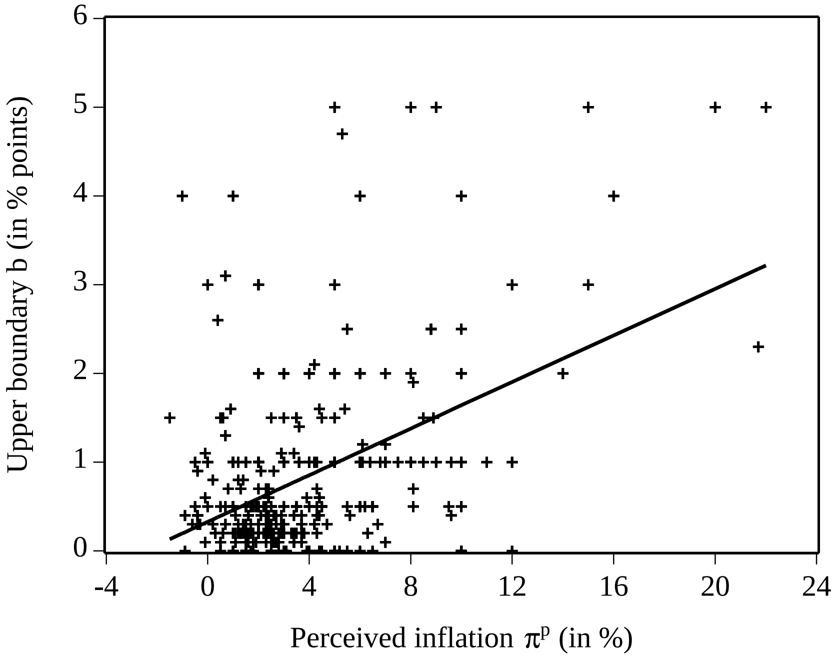


Figure 1.4: Estimation of the upper boundary

a fall, because it is not clear whether the difference in the absolute values of the constants and the slope coefficients are significant. We can test for asymmetries in a more formal way by running a pooled regression of the following type:

$$|x| = \phi_0 + \phi_1 d + \phi_3 \pi^p + \phi_4 \pi^p d + \varepsilon_x. \tag{1.12}$$

The vector x contains the values of a and b; d is a dummy variable that is equal to one if $x = a$ and zero otherwise; ε_x is the error term. If the estimated coefficients shown in table 1.5 are statistically identical, then both coefficients ϕ_1 and ϕ_4 should not be significant. The p-values we obtained are 0.03 for ϕ_1 and 0.36 for ϕ_4. Thus, the lower boundary is significantly smaller than the upper boundary by a constant value of about $0.3288 - 0.1388 \approx 0.2$ percentage points. However, the slope coefficients in equations (1.10) and (1.11) are not statistically different. This can be seen from the high p-value associated with ϕ_4. Thus, we conclude that the asymmetrical behavior does not change for different π^p and the difference between the absolute values of the boundaries a and b stays constant. It follows that respondents seem to react more sensitively to an expected fall of the inflation rate than to a rise.

Another point that can be made here is that the linear fit of the OLS regressions in table 1.5 might not give a good approximation of the just noticeable difference.

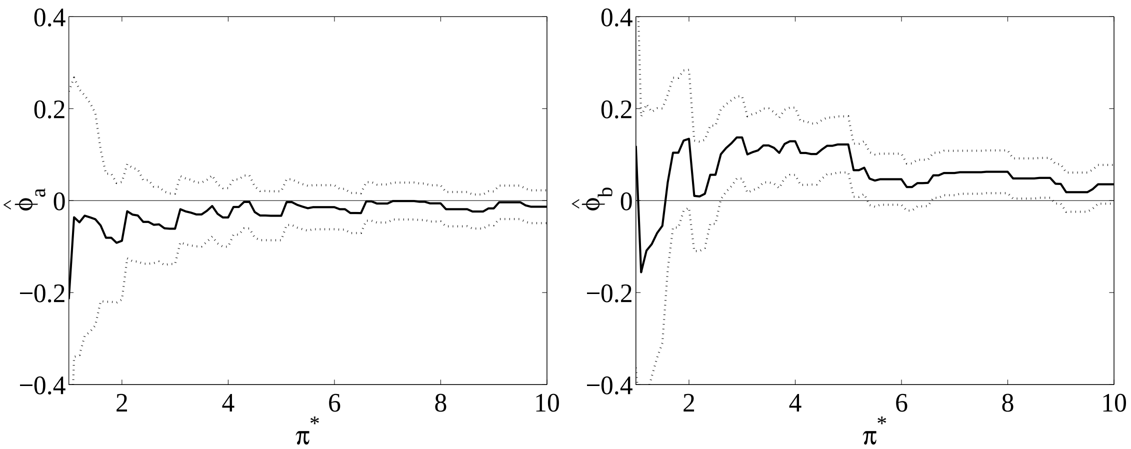


Figure 1.5: Linearity of the Weber-Fechner law

It may be the case that respondents from countries with a low perceived inflation rate may form a different attitude towards expected changes of the inflation rate than respondents from countries experiencing high inflation. This would imply that the slope coefficient varies with the perceived inflation rate, which would give rise to non-linearities. To elaborate on this, we run a series of regressions of the following type:

$$a = \delta_0 + \delta_1 \pi^p + \phi_a d_a \pi^p + \varepsilon_a \tag{1.13}$$

$$b = \gamma_0 + \gamma_1 \pi^p + \phi_b d_b \pi^p + \varepsilon_b, \tag{1.14}$$

where d_a (d_b) is a dummy vector in which the i-th row $d_{ai} = 1$ ($d_{bi} = 1$) if $\pi_i^p \geq \pi^*$. The index i refers to an individual respondent. π^* rises from 1% to 10% in steps of 0.1 so that the division line between countries with low inflation and countries with high inflation is variable. For each regression we record $\hat{\phi}_a$ and $\hat{\phi}_b$ as well as the respective 95% confidence bands. The results are summarized in figure 1.5. The left panel shows that $\hat{\phi}_a$ is not significantly different from zero irrespective of the value of π^*. This clearly indicates that a linear fit is appropriate for the estimation of the lower boundary. The right panel shows the results for the upper boundary. Here, $\hat{\phi}_b$ turns out to be positive and significant for values of π^* between 2.6% and 5.3% and for values of π^* between 6.6% and 8.8%. In order to use these results for the calculation of the upper boundary, a decision has to be made on the value of π^*. Using the highest R^2 as criterion, $\pi^* = 5\%$ which results in the estimates

	$\hat{\gamma}_0$	$\hat{\gamma}_1$	$\hat{\phi}_b$	R^2
Upper boundary b	0.5833⋆	−0.0046	0.1221⋆	0.2345
	(0.0920)	(0.0371)	(0.0311)	

Table 1.6: Linearity of the Weber-Fechner law ($\pi^* = 5\%$)

shown in table 1.6. Note that standard errors are given in parentheses. A ⋆ indicates significance at the 5% level.

As $\hat{\gamma}_1$ is insignificant, the results imply that for values of $\pi_i^p < 5\%$ the upper boundary of the just noticeable difference is constant and does not rise with the perceived inflation rate. By contrast, for values of $\pi_i^p \geq 5\%$ the slope coefficient $\hat{\phi}_b$ becomes positive and significant so that the upper boundary increases with the perceived inflation rate. Interestingly, $\hat{\phi}_b$ is statistically indifferent from $\hat{\gamma}_1$ of the baseline regression (1.11). Although we could conclude that there is some evidence of a non-linear relationship between the upper boundary of the just noticeable difference and the perceived inflation rate, we decided to use the linear baseline regression (1.11) for the conversion of qualitative expectations into quantitative measures below in section 1.4.4. This basically has two reasons. First, equation (1.14) is just one formulation of a non-linear regression model among many possible models. Thus, further research on the theoretical foundations of the evolution of the just noticeable difference would be needed in order to specify a model that allows for deviations from the linear Weber-Fechner law. Second, the improvement in terms of R^2's when using equation (1.14) instead of equation (1.11) is only 1.6% and in our view too small in order to justify a more complicated behavioral model.

Apart from testing for non-linearities, the answers to our additional question can be used to investigate whether or not the just noticeable difference varies across countries. Equations (1.10) and (1.11) are estimated from a cross section and it is assumed that the only source of variation in the just noticeable difference is the perceived inflation rate. There may however be one good reason to suppose that the estimated results might suffer from an omitted variable bias because of unobserved heterogeneity. If agents were used to live in an environment with a high average level of inflation over a long period of time, they might have a lower sensibility towards changes in inflation than those who have never been faced with high inflation rates. Even though in 2004 only 10 percent of the countries in our sample experienced inflation rates of above 10 percent (and below 22 percent), this share has been much higher in the preceding decades. In 48 countries the average inflation rate between 1973 and 1990 was above 10 percent, and in 18 countries inflation even exceeded 50 percent.[12] In order to test whether the inflationary history of a country

[12]For the Eastern European countries and the countries of the former Soviet Union the average

has any specific impact on the just noticeable difference of the respondents, we additionally controlled for it and re-estimated equations (1.10) and (1.11). As it is a stylized fact that the variability of inflation is positively correlated with the average level of inflation, the inflationary history of each country was approximated by the average standard deviation of the annual inflation rates between 1973 and 1990. The results including standard errors in parentheses are shown in table 1.7. A * indicates significance at the 5% level.

	$\hat{\delta}_0$	$\hat{\delta}_1$	$\hat{\delta}_2$	R^2
Lower boundary a	$-0.1320^\star$ (0.0561)	$-0.1375^\star$ (0.0115)	$-0.00047^\star$ (0.00013)	0.3484
	$\hat{\gamma}_0$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	R^2
Upper boundary b	$0.3227^\star$ (0.0658)	$0.1204^\star$ (0.0139)	$0.00040^\star$ (0.00013)	0.2232

Table 1.7: Controlling for country-specific inflationary history

The estimated coefficients for the impact of the inflationary history $\hat{\delta}_2$ and $\hat{\gamma}_2$ are significant and show the expected sign. Therefore, the just noticeable difference has to be adjusted for an additional country-specific constant term that decreases the lower boundary and that increases the upper boundary. Compared with the baseline estimates shown in table 1.5 the goodness-of-fit of both regressions improves by roughly 2 percentage points. The magnitude of the country effect, however, is so small that in the end we decided not to consider it when converting the qualitative expectations into quantitative measures below in section 1.4.4. For the UK, for instance, where inflation exhibited the highest standard deviation in 1973-90 (5.65%) among the countries investigated in section 1.3, the lower (upper) boundary a (b) would have to be corrected by -0.0026 (0.0022). Note also that the estimated parameters of the Weber-Fechner law ($\hat{\delta}_0$, $\hat{\delta}_1$, $\hat{\gamma}_0$ and $\hat{\gamma}_1$) are statistically indifferent from the baseline estimates shown in table 1.5 as they are within the 95% confidence band.

1.4.3 Identifying the Basis of the Expectation Formation Process

From part a) of our additional question we can finally infer the inflation rate which was perceived by the respondents at the time they filled in the *CESifo WES* questionnaire (between July 05 and July 15, 2004). On the basis of these answers, we are

inflation rate between 1991 and 1995 was used to calculate these figures. The source of the annual inflation rates are the World Bank's *World Development Indicators*.

able to detect the average information lag of the respondents and, hence, the basis on which they form expectations. As the variation of the answers for each country is large, we use measures of average deviation. The root mean squared error (RMSE) and the mean absolute error (MAE) can be calculated such that they measure the deviation of the inflation rate reported in response to question a) from the reference inflation rate prevailing in the current and previous quarters and months of 2004. For the non-OECD countries in our sample the data was taken from the *International Financial Statistics* of the IMF. The main difference between the RMSE and the MAE is that the RMSE puts more weight on deviations that are large.

	MAE		RMSE
M6 2004	1.17	M6 2004	1.84
M5 2004	1.14	M5 2004	1.83
M4 2004	1.22	M4 2004	1.94
M3 2004	1.35	M3 2004	2.03
M2 2004	1.40	M2 2004	2.20
M1 2004	1.44	M1 2004	2.70
Q2 2004	1.06	Q2 2004	1.74
Q1 2004	1.29	Q1 2004	2.01
Q4 2003	1.58	Q4 2003	2.43

Table 1.8: Identifying the basis for inflation expectations

The calculation of the errors, which are presented in table 1.8, shows that the smallest error is calculated for the second quarter of 2004 for both measures when the analysis is done on a quarterly basis. On a monthly basis, the results are qualitatively the same. The values for the RMSE and the MAE suggest that the smallest error emerges if May 2004 is taken as a reference. Nevertheless, the smallest MAE is indicated for the second quarter of 2004 and the RMSE for May 2004 is only slightly below the one calculated for the second quarter 2004.

Interestingly, both, the MAE and the RMSE are smaller for the quarterly series. From this and from the fact that inflation expectations of the *CESifo WES* are 6-months-ahead inflation expectations which are queried every three months in the first two weeks of January, April, July and October, we conclude that a quarterly perspective seems most appropriate. The information set that is available to the survey respondents at the time they fill in the questionnaire is the past quarter (that is the first quarter for the questionnaires returned at the beginning of April, the second quarter for the questionnaires returned at the beginning of July, and so on). Thus, the July survey produces inflation expectations $E_t\pi_{t+2}$, where t refers to the second quarter and $t+2$ to the fourth quarter.

1.4.4 Application to the *CESifo WES*

Given the basis of the expectation formation process and the survey-based estimates of the just noticeable difference, it is now possible to convert qualitative inflation expectations for every country considered by the *CESifo WES* into quantitative measures of expectations using equation (1.3). The upper and lower boundary of the indifference interval are calculated according to

$$a_t = \hat{\delta}_0 + \hat{\delta}_1 \pi_t \quad \text{and} \quad b_t = \hat{\gamma}_0 + \hat{\gamma}_1 \pi_t$$

with $\hat{\delta}_0$, $\hat{\delta}_1$, $\hat{\gamma}_0$ and $\hat{\gamma}_1$ taken from table 1.5. Thus, we assume that these behavioral parameters are equal across countries and across agents, that they are asymmetric for the upper and lower boundary, and that the just noticeable difference only varies proportionally with the basis of the expectation formation process π_t.

The results of the conversion together with the balance statistic are shown in figures 1.6 and 1.7. Note that the outliers resulting from small sample sizes still occur in France (2001:1), Italy (1996:4) and the UK (1991:4), despite our adjustments described in appendix 1.A.

1.4.5 Unbiasedness of Expectations

With the quantified inflation expectations at hand, it is now possible to test whether or not expectations are unbiased predictors of future realizations. While the traditional conversion methods impose the unbiasedness of expectations, our approach allows to test for this property. The reason why we are questioning this assumption is the mixed evidence reported in the literature. Many papers that have examined quantitative survey measures of inflation expectations (which were directly queried like in the US-based Livingston or Michigan survey and which have not been converted from qualitative data) have concluded that these expectations are biased forecasts of inflation (see Roberts (1997) and the papers cited there).

The unbiasedness of expectations constitutes a necessary condition for rational expectations in the sense of Muth (1961). Unbiasedness implies that the forecast error should, on average, be equal to zero. For a forecast horizon of two quarters this hypothesis is typically tested by estimating the following equation:

$$\pi_t - E_{t-2}\pi_t = c + u_t. \tag{1.15}$$

If the null hypothesis that $c = 0$ can be rejected at reasonable levels, we conclude that expectations were indeed biased. The results of this regression for all conversion methods applied in this paper are given in table 1.9. The p-values for the t-tests which have been calculated using Newey-West standard errors to correct for

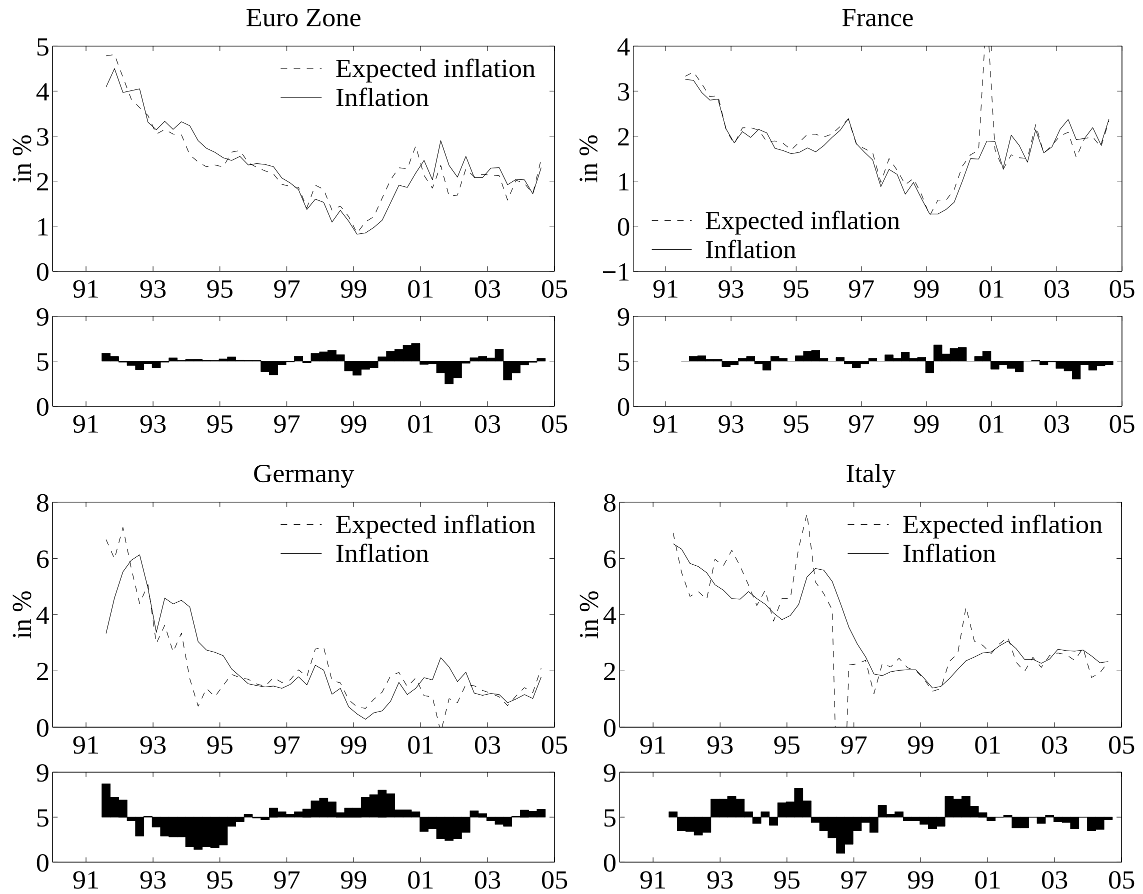


Figure 1.6: Quantification with the survey-based method

overlapping forecast errors, are reported in brackets. A * indicates significance at the 5% level.

The unbiasedness tests for the inflation expectations obtained from the Carlson-Parkin method and the Weber-Fechner method show that all constant terms are close to zero and insignificant at the 5% level. Of course, this result does not come as a surprise as the unbiasedness is a crucial assumption for each method. The unbiasedness tests for the regression approach are presented in the next two columns. As far as OLS results are concerned, they yield perfectly unbiased expectations. This is due to the fact that the expectation error is identical to the estimated error $\hat{\varepsilon}_t$ of the OLS regression.[13] The time-varying parameters approach yields very similar results

[13]See equation (1.8) and footnote 9 in section 1.3.2. In the cases of France and Italy we get a point estimate that deviates from zero because the dummy variable that we included for the estimation of the boundaries is not taken into account when calculating expectations.

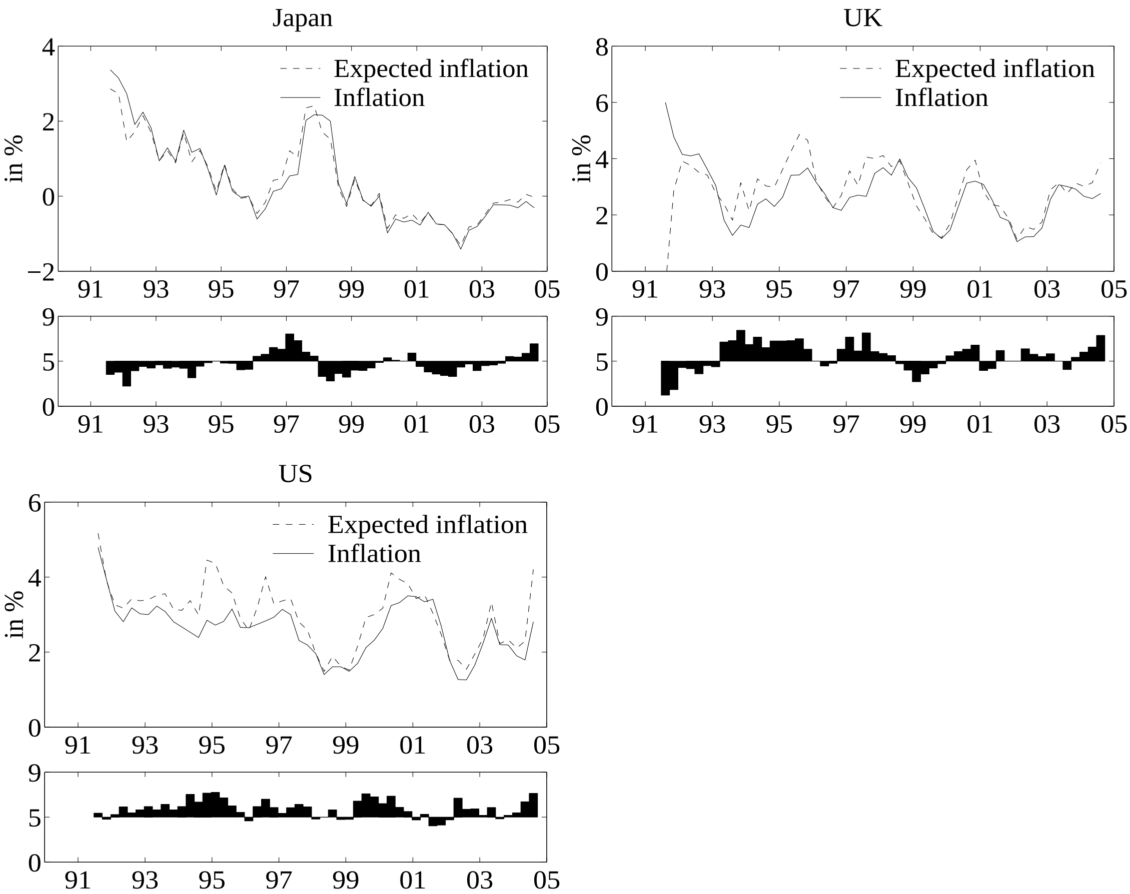


Figure 1.7: Quantification with the survey-based method ctd.

with all the estimates being close to zero and insignificant at the 5% level. Note, that the smoothed component of the boundaries is used to calculate expectations.

In the last column of table 1.9 we present the results for the survey-based method. In fact, it turns out that for two out of the seven countries inflation expectations were biased during the period 1991:2 to 2004:2. Here, Italy and the US do not fulfill the necessary condition for rationality.[14] This finding is very much in line with our conjecture in section 1.3.1 where we argued that the negative just noticeable difference for these countries that has been derived from the Carlson-Parkin method is a result of 'bad' expectations or, to put it more concretely, of expectations that were biased upwards throughout the period of disinflation in the

[14]Using the Livingston Survey of Professional Forecasters (which queries quantitative inflation expectations) Adam and Padula (2003) find that expectations in the US were indeed biased during the nineties.

Country	Carlson-Parkin	Weber-Fechner	Regression OLS	Regression TVP	Survey-based
Euro Zone:	−0.01 [0.92]	−0.06 [0.29]	0.00 [1.00]	−0.01 [0.62]	−0.06 [0.42]
France:	0.03 [0.86]	−0.01 [0.86]	−0.04 [0.54]	−0.04 [0.31]	−0.08 [0.34]
France:$^\diamond$	0.01 [0.95]	0.02 [0.86]	0.00 [1.00]	−0.01 [0.75]	
Germany:	−0.02 [0.86]	−0.07 [0.58]	0.00 [1.00]	−0.01 [0.83]	0.05 [0.78]
Italy:	0.03 [0.99]	−0.10 [0.82]	0.07 [0.38]	0.01 [0.82]	−0.22* [0.04]
Italy:$^\diamond$	0.00 [0.99]	−0.12 [0.18]	0.00 [1.00]	−0.01 [0.73]	
Japan:	0.03 [0.73]	−0.02 [0.85]	0.00 [1.00]	−0.00 [1.00]	−0.13 [0.13]
UK:	0.11 [0.63]	0.02 [0.90]	0.00 [1.00]	0.00 [0.94]	−0.22 [0.24]
UK:$^\diamond$	0.12 [0.72]	0.73 [0.42]	0.00 [1.00]	0.00 [0.98]	
US:	0.06 [0.66]	−0.02 [0.90]	0.00 [1.00]	−0.00 [1.00]	−0.45* [0.00]

Note: For the countries in which outliers occurred due to the conversion of inflation expectations from qualitative into quantitative data the analysis of the forecast error was done with and without dummy variable. In case the dummy variable was included, we used the indifference bands that were calculated excluding the outliers. A $\diamond$ marks the estimations where no dummy variable was included.

Table 1.9: Unbiasedness tests

beginning of the 1990s. The negative sign of the constant confirms this conjecture. In general, table 1.9 reveals that with the traditional methods we are not capable of testing the unbiasedness of expectations or do any inference on rational expectations.

1.5 Conclusion

This paper presents a new methodology for the determination of the just noticeable difference, which is required for the quantification of qualitative survey data. Traditional conversion methods, such as the probability approach of Carlson and Parkin (1975), the regression method of Pesaran (1984) or the time-varying parameters model of Seitz (1988), require very restrictive assumptions concerning the properties of the just noticeable difference and the expectation formation process of survey respondents. The novelty of the present paper is that we convert qualitative survey responses into quantitative measures for inflation expectations without having to rely on these assumptions. In contrast to the three traditional methods we do not implicitly derive the just noticeable difference from the qualitative survey responses and from the statistical properties of the reference time-series, but from a special question in the July 2004 *CESifo WES*, in which we directly query the repondents' boundaries of the indifference interval for a given current inflation rate. This new methodology, which we labeled survey-based approach, was then applied to expectations about the future development of inflation which are included in the *CESifo WES*.

The major advantage of our approach is that we can explicitly test whether or not the assumptions made in the traditional conversion methods are valid. Specifically, we addressed the following issues:

- are the boundaries symmetric and constant over time;
- if not, what are the determinants of the boundaries;
- do the boundaries vary across countries;
- are expectations unbiased?

Concerning the first two issues, our main results are that boundaries are asymmetric and time-varying. Respondents seem to react more sensitively to an expected fall of the inflation rate than to a rise. Moreover, the boundaries turned out to be an increasing function of the perceived current rate of inflation. With the Weber–Fechner law we delivered a theoretic rationale for this relationship. Concerning the third issue, we found that there are country-specific effects which are related to the country's inflationary history, but these effects are so small that we decided not to consider them for the conversion. Concerning the final issue we showed that the unbiasedness assumption that is made in all traditional conversion method holds for a majority of the countries in our sample, but not for all.

Apart from the relaxation of some crucial assumptions underlying the traditional conversion methods, a more practical advantage of the survey-based method is that the resulting time series for inflation expectations are not subject to revisions. While in the traditional methods the just noticeable difference is recalculated with every additional data point, in our approach the boundaries are exogenous to qualitative expectations and only vary with the level of the current rate of inflation.

The problems related to the assumption of normally distributed survey responses remain unsolved by our approach. Like the traditional conversion methods, the survey-based method uses the computed boundaries to divide the probability density function of the normal distribution into three sub-areas: expectations of a lower, a constant and a higher future inflation rate. Problems emerge when there are no survey participants in one of the categories. This situation appears quite often in an expert survey such as the *CESifo WES* with a limited number of participants; so that in the present paper, we only considered countries for which a critical number of respondents was exceeded.

Appendix

1.A Corrections of the Microdata

A major shortcoming of the Carlson-Parkin method is the underlying assumption that aggregate distribution of responses is approximated by a normal distribution. There are three cases in which the calculation of a quantitative measure of inflation expectations according to equation (1.3) becomes impossible:

- no respondent is within the category *UP* or *DOWN*,
- all respondents share the same opinion,
- no respondent is within the category *SAME*.

First, if $UP_t = 0$ the value of $\Phi^{-1}(1 - UP_t)$ approaches infinity, whereas the value of $\Phi^{-1}(DO_t)$ approaches minus infinity whenever $DO_t = 0$. If such a case occurred, we corrected for that by adding $1/(2n + 1)$ to the category that is equal to zero, with n being the number of respondents at time t, and by subtracting this value from the opposite category. This can be justified by the fact that the answers of the survey only approximate the basic population. With this correction we do not fundamentally change the survey result as the number of respondents stays the same when the corrected figures are rounded to nearest integer[15].

Second, if either $UP_t = 1$, or $SAME_t = 1$, or $DO_t = 1$, we subtracted $1/(2n+1)$ from the respective category. In contrast to the first case, the remaining two categories are increased by only $1/[2(2n + 1)]$ so as to obtain a non-zero fraction in every category.

[15] Take the following outcome of the survey as an example: $UP_t = 0$, $SAME_t = 0.5$, $DO_t = 0.5$ and $n = 10$. Applying the correction mechanism yields the following adjusted fractions: $UP_t = 0.048$, $SAME_t = 0.5$ and $DO_t = 0.452$. For a number of ten respondents this gives 0.48 persons expecting a rise in inflation which is equal to zero when rounded to nearest whole number, and 4.52 persons expecting a fall in inflation which can be rounded to 5.

Third, if we obtain a value of $SAME_t = 0$, the denominator in formula (1.3) is zero. To avoid this problem, we subtracted $1/[2(2n+1)]$ from the UP_t and DO_t fractions and added $1/(2n+1)$ to the $SAME_t$ fraction.

1.B Estimates of the Just Noticeable Difference from the Time-Varying Parameters Method

In section 1.3.2 we calculated the permanent component of the just noticeable difference by implementation of the time-varying parameters model. For illustration we depict these results in figure 1.8. The solid line shows the upper boundary b and the dashed line shows the lower boundary a. A $^{\diamondsuit}$ marks the cases where the outliers are included into the regression.

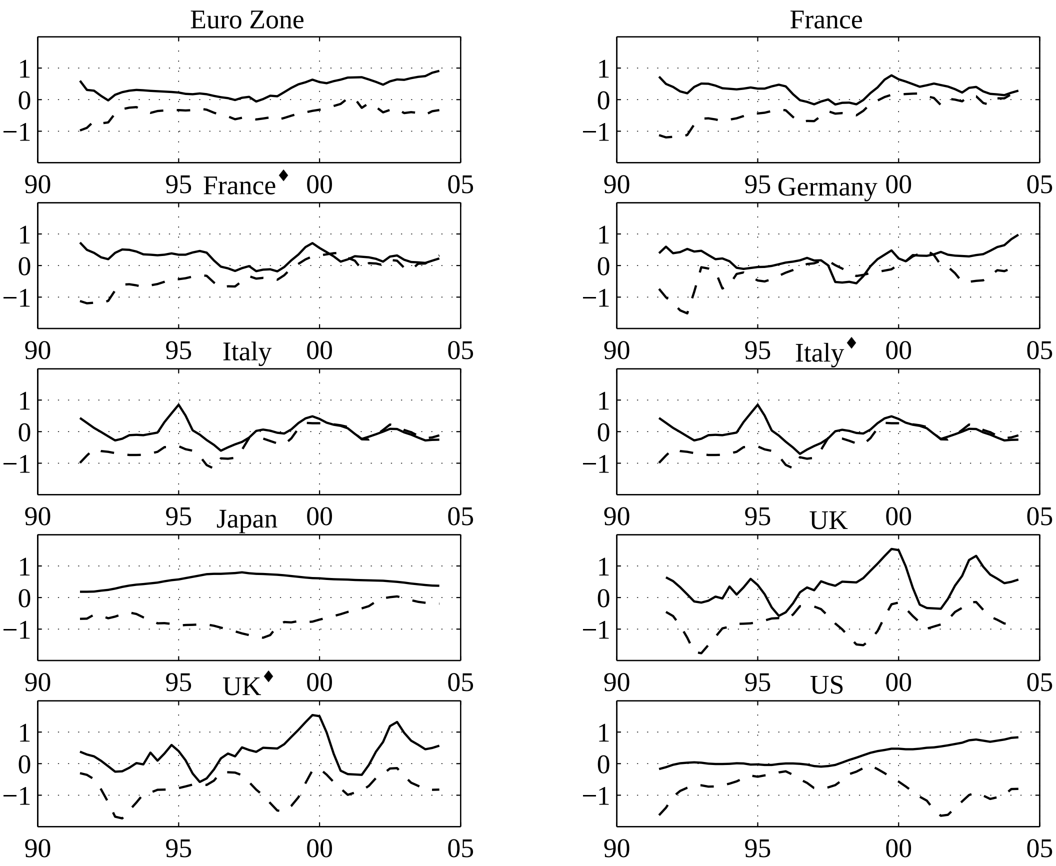


Figure 1.8: Just noticeable differences estimated with the time-varying parameters method

Chapter 2

The New Keynesian Phillips Curve and the Role of Expectations: Evidence from the CESifo World Economic Survey

Abstract

We provide evidence on the fit of the hybrid New Keynesian Phillips curve for selected Euro zone countries, the US and the UK. Instead of imposing rational expectations and estimating the Phillips curve by the Generalized Method of Moments, we use direct measures of inflation expectations from the CESifo World Economic Survey. Our main findings are as follows: (i) The use of survey data gives empirical results, which are more reliable than those obtained from the GMM approach. (ii) The purely forward-looking Phillips curve can be rejected in favor of the hybrid New Keynesian Phillips curve. (iii) The estimated coefficients on past inflation are higher when using survey expectations than when using the rational expectations GMM approach. (iv) It remains unclear whether real unit labor costs or a measure of the output gap should be used as a proxy for real marginal costs. (v) Theory-based restrictions lead to an improvement of the empirical results.

2.1 Introduction

The relationship between inflation and real variables is of crucial importance for understanding the effects of monetary policy on inflation. In recent years, some kind of consensus has emerged, generally referred to as New Keynesian macroeconomics, that integrates Keynesian elements (imperfect competition, nominal rigidities) into a dynamic general equilibrium framework traditionally used in the Real Business Cycle literature. The nature of inflation dynamics is arguably the most distinctive feature of the New Keynesian paradigm. It is captured by the so-called New Keynesian Phillips curve which is based on Calvo's (1983) model of staggered price setting and which expresses current inflation as a function of expected future inflation and a measure of firms' real marginal costs. While theoretically appealing, a number of authors (e.g. Fuhrer and Moore, 1995) criticized this version of the Phillips curve since the implied 'jump' behavior of inflation was completely at odds with the hump-shaped behavior that can be observed in VAR analyses. As a consequence, Galí and Gertler (1999) extended Calvo's theoretical framework to the so-called hybrid New Keynesian Phillips curve (HNKPC) by allowing for a fraction of firms that set prices according to a backward-looking rule-of-thumb.

The empirical findings are encouraging for the New Keynesian Phillips curve. Empirical work mainly centers around the question of which variable to use for measuring real activity and whether backward-looking behavior is relevant. Concerning the first question, theory tells us that real marginal costs are the driving force underlying changes in inflation. In a recent survey article Galí (2003) emphasizes that empirical results are promising when the New Keynesian Phillips curve is estimated in a way consistent with theory, implying that labor income share is used instead of detrended GDP as a proxy for real marginal costs. Concerning the second question he continues: "Although backward-looking behavior is often statistically significant, it appears to have limited quantitative importance. In other words, although the baseline purely forward-looking model is rejected on statistical grounds, it is still likely to be a reasonable first approximation to the inflation dynamics of both Europe and the United States." (ibid., p. 162).

The standard econometric tool for estimating the New Keynesian Phillips curve is the Instrumental Variables technique or, more generally, the Generalized Method of Moments (GMM). Expectations about future variables are replaced by their ex-post realizations, and expectational errors are assumed to be uncorrelated with all variables in the information set of agents available at the time expectations are formed. In other words, expectations are assumed to be rational. There is, however, an ongoing debate in the recent literature about the appropriateness of the GMM technique. As mentioned by Mavroeidis (2005) and Rudd and Whelan (2005), GMM estimates may overstate the degree of forward-looking behavior if the Phillips curve model is mis-specified.

The contribution of this paper is twofold. On a theoretical level we derive the HNKPC under the assumption that firms have subjective expectations that may be non-rational. Available evidence from surveys suggests that inflation expectations are often biased and inefficient predictors of future inflation, thereby questioning the assumption of rationality (see Roberts, 1997, and the papers cited there). We extend the theoretical framework developed by Adam and Padula (2003) by allowing for the existence of both, forward-looking and backward-looking firms. On an empirical level we follow Roberts (1997) and Adam and Padula (2003) and estimate the Phillips curve for selected Euro zone countries, the US and the UK by using direct measures of inflation expectations, instead of imposing rational expectations and estimating the Phillips curve by GMM. The data source is the CESifo World Economic Survey which quarterly polls economic experts about their expected future development of inflation. Our main findings are as follows: (i) The use of survey data gives empirical results, which are more reliable than those obtained from the rational expectations GMM approach. We show that OLS parameters are stable over time and endogeneity of regressors can be rejected. (ii) The purely forward-looking Phillips curve can be rejected in favor of the hybrid New Keynesian Phillips curve. (iii) The estimated coefficients on past inflation are higher when using survey expectations than when using the rational expectations GMM approach. (iv) It remains unclear whether real unit labor costs or a measure of the output gap should be used as a proxy for real marginal costs. (v) Theory-based restrictions lead to an improvement of the empirical results.

The paper is organized as follows. In section 2.2 we present the standard version of the HNKPC that results from a rational expectations approach and we modify it in a way that accounts for subjective and potentially non-rational expectations of firms. Section 2.3 gives an overview of the data. The main focus is on the presentation of the inflation expectations from the CESifo WES, but we also briefly discuss the variables used as proxies for real marginal costs. Our estimation results and a comparison with other empirical work (mainly using the rational expectations approach) are presented in sections 2.4 and 2.5. Finally, section 2.6 summarizes the main results and concludes.

2.2 The Hybrid New Keynesian Phillips Curve

2.2.1 Rational Expectations

The version of the HNKPC that is mostly used in the literature has been introduced by Galí and Gertler (1999) and extended by Galí, Gertler, and López-Salido (2001).

It is based on Calvo's (1983) staggered price setting framework in which each firm has a probability $1-\theta$ of being able to reset its price in any given period, independently of the time elapsed since the most recent price adjustment. In contrast to Calvo (1983), however, they assume that of those firms being able to adjust prices in a given period, there is only a fraction of firms $1-\omega$ that sets prices optimally in a forward-looking manner. The remaining part uses a rule-of-thumb that simply augments last period's average reset price by the inflation rate prevailing in that period. It can then be shown that the HNKPC is given by

$$\pi_t = \gamma_f E_t[\pi_{t+1}] + \gamma_b \pi_{t-1} + \lambda mc_t \tag{2.1}$$

where π_t denotes the inflation rate, $E[\cdot]$ the rational expectations operator, and mc_t the logarithm of real marginal costs, and where the coefficients can be expressed in terms of the structural parameters

$$\gamma_f = \frac{\beta\theta}{\theta + \omega[1-\theta(1-\beta)]},$$

$$\gamma_b = \frac{\omega}{\theta + \omega[1-\theta(1-\beta)]},$$

$$\lambda = \frac{(1-\omega)(1-\theta)(1-\beta\theta)}{\theta + \omega[1-\theta(1-\beta)]}.$$

β is the discount factor of the firms' intertemporal maximization problem. An important assumption underlying the derivation of the structural parameters is that firms operate under monopolistic competition with a Cobb-Douglas production technology and constant returns to scale. If returns to scale are decreasing, Galí, Gertler, and López-Salido (2001) showed that λ additionally becomes a function of the labor elasticity of production and the price elasticity of demand.

This very general formulation of the Phillips curve comprises two special cases. First, the discount factor β can be restricted to unity, $\gamma_f + \gamma_b = 1$, which implies that in the long-run the Phillips curve is vertical. Second, when $\omega = 0$ all firms set their prices optimally and the model converges to the purely forward-looking New Keynesian Phillips curve ($\gamma_f = \beta$, $\gamma_b = 0$, $\lambda = [(1-\theta)(1-\beta\theta)]/\theta$).

2.2.2 Subjective Expectations

As in the previous section we distinguish between two groups of firms: forward-looking firms which set prices according to an intertemporal optimization procedure, and backward-looking firms which set prices according to a simple rule-of-thumb. The main difference to the previous section is the way forward-looking firms form their expectations. Instead of imposing rational expectations (i.e. all firms form expectations homogenously, using the same model and the same information set), we allow for subjective expectations of each single forward-looking firm, which may be rational or not and which may be heterogeneous across firms.

In the following we will derive the HNKPC under the assumption that firms form subjective expectations. We will extend the theoretical framework of Adam and Padula (2003) by explicitly introducing backward-looking firms. In contrast to their paper which describes the price-setting behavior of firms from the point of view of professional forecasters, we assume that the source of potential non-rationalities in expectations are the firms themselves. This has the advantage that we can continue to distinguish between two types of firms as in the case of rational expectations.

In accordance with the rational expectations approach, the starting point is Calvo's (1983) staggered price setting framework, which defines the log of the aggregate price level p_t as

$$p_t = (1-\theta)p_t^* + \theta p_{t-1}, \tag{2.2}$$

where p_t^* is the average reset price and $1-\theta$ the probability that firms reset prices. The average reset price is a weighted sum of the average price set by forward-looking firms and the average price set by backward-looking firms

$$p_t^* = (1-\omega)\frac{1}{I}\sum_{i=1}^{I} p_t^{f,i} + \omega\frac{1}{J}\sum_{i=1}^{J} p_t^{b,i}, \tag{2.3}$$

where I (J) is the number of forward-looking (backward-looking) firms, ω the fraction of backward-looking firms ($\omega = J/(I+J)$), and $p_t^{b,i}$ ($p_t^{f,i}$) the price set by the backward-looking (forward-looking) firm i. All firms which set prices in a backward-looking manner, follow an identical rule-of-thumb according to which last period's average reset price is simply corrected by lagged inflation. Forming the average of all backward-looking firms gives

$$p_t^b = \frac{1}{J}\sum_{i=1}^{J} p_t^{b,i} = p_{t-1}^* + \pi_{t-1}. \tag{2.4}$$

Firms which behave in a forward-looking manner, maximize expected discounted profits given technology, factor prices and the constraint on price adjustment (defined by $1-\theta$) which results in the following log-linear rule:

$$\begin{aligned} p_t^{f,i} &= (1-\beta\theta)F_t^i[\sum_{k=0}^{\infty}(\beta\theta)^k(mc_{t+k}+p_{t+k})] = \\ &= (1-\beta\theta)(mc_t+p_t) + \beta\theta F_t^i[p_{t+1}^{f,i}], \end{aligned} \tag{2.5}$$

where $F_t^i[\cdot]$ denotes the subjective expectations operator of firm i.[1] While individual firms produce differentiated products under monopolistic competition, they are all assumed to have the same Cobb-Douglas production technology and to face demand curves with constant and equal demand elasticities. The crucial problem now is the aggregation of individual prices set by forward-looking firms. Following Adam and Padula (2003) we assume that firm i forms expectations about other firms' optimum prices and aggregates them to the average forward-looking price:

$$F_t^i[p_{t+1}^f] = F_t^i[\frac{1}{I}\sum_{h=1}^{I} p_{t+1}^{f,h}]. \tag{2.6}$$

Defining the average current forward-looking price by

$$p_t^f = \frac{1}{I}\sum_{h=1}^{I} p_t^{f,h} \tag{2.7}$$

and assuming that the 'law of iterated expectations' holds, which implies that agents do not expect that current forecasts of future variables z will be revised in a particular direction in the next period, i.e.:

$$F_t^i[F_{t+1}^h[z_{t+s}] - F_t^h[z_{t+s}]] = 0 \quad \forall\ i, h, s > 0, \tag{2.8}$$

Adam and Padula (2003) show that equation (2.6) can be expressed as

$$F_t^i[\pi_{t+1}^f] = (1-\beta\theta)(F_t^i[p_{t+1}^f] - mc_t - p_t), \tag{2.9}$$

where $\pi_{t+1}^f = p_{t+1}^f - p_t^f$. In order to get this equation they take the difference between equation (2.6) and (2.7), replaced $p_t^{f,h}$ with the first expression of equation (2.5) and applied the law of iterated expectations (see appendix 2.A).

Combining equations (2.2), (2.3) and (2.4) gives a relationship between p_t^f and p_t (see appendix 2.B),

$$p_t^f = \frac{p_t + (\theta\omega - 2\omega - \theta)p_{t-1} + \omega p_{t-2}}{(1-\theta)(1-\omega)}, \tag{2.10}$$

which can be shifted one period forward by applying the $F_t^i[\cdot]$ operator:

$$F_t^i[p_{t+1}^f] = \frac{F_t^i[p_{t+1}] + (\theta\omega - 2\omega - \theta)p_t + \omega p_{t-1}}{(1-\theta)(1-\omega)}. \tag{2.11}$$

[1] Apart from the $F_t^i[\cdot]$ operator, equation (2.5) is identical with the optimum pricing rule under rational expectations. For a derivation see Galí and Gertler (1999) and Galí, Gertler, and López-Salido (2001).

Inserting equations (2.10) and (2.11) into equation (2.9) and aggregating over all subjective expectations, $\bar{F}_t[\cdot] = (1/I)\sum_{i=1}^{I} F_t^i[\cdot]$, finally gives the HNKPC based on average subjective expectations (see appendix 2.C),

$$\pi_t = \gamma_f \bar{F}_t[\pi_{t+1}] + \gamma_b \pi_{t-1} + \lambda mc_t, \tag{2.12}$$

where $\pi_t = p_t - p_{t-1}$. Note that equation (2.12) is identical with the specification derived under rational expectations, except for the way expectations are formed.

2.3 Data Description

2.3.1 Inflation Expectations from the CESifo World Economic Survey

Subjective inflation expectations are taken from the CESifo World Economic Survey (WES) which assesses trends in the world economy by polling transnational as well as national organizations worldwide about economic developments. It is conducted in co-operation of Ifo Institute for Economic Research and the International Chamber of Commerce (ICC) in Paris. The questionnaire of the WES, which is distributed every quarter (January, April, July and October) and which was first conducted in March 1983, asks participants to give their assessment of the general economic situation and expectations regarding important macroeconomic indicators of the country they inhabit. Currently, the WES asks about 1100 experts in 90 countries.

A question on the expected inflation rate, which is in the focus of the present paper, was only included since July 1991. Survey participants are asked to give their expectations on the inflation rate by the end of the next six months. They indicate UP for an expected rise in the inflation rate, SAME for no change in the inflation rate and DOWN for an expected fall in the inflation rate by the end of the next six months. The questionnaire therefore reveals qualitative information on the participants' expectations, which we transformed into a time series of expected cardinal inflation rates by applying a variant of the probability approach of Carlson and Parkin (1975).

The probability approach assumes that the median expectations of the respondents sampled are normally distributed and that respondents report expected changes only if these changes are above or below a certain threshold. To derive an estimate of these thresholds, which define the so-called indifference interval, it is further assumed that they are symmetric around zero and constant across time and individuals. The indifference interval is then calculated in such a way that throughout the observation period the constructed expected inflation rates are on average equal to the actual inflation rates. This implies a priori that inflation expectations are unbiased predictors of future inflation, which is a necessary condition for rationality. In Henzel and Wollmershäuser (2005) we presented a new methodology for the determination of the thresholds of the indifference interval that avoids this assumption. Instead of deriving these thresholds by imposing that the constructed expected inflation rates are on average equal to the actual inflation rates, we directly queried them from the survey respondents by a special question in the CESifo WES. The main results are that the thresholds are asymmetric and time-varying. Respondents seem to react more sensitively to an expected fall of the inflation rate than to a rise. Moreover the thresholds turned out to be an increasing function of the perceived current rate of inflation.[2]

The converted inflation expectations and the actual inflation rate for Germany, France, Italy, the Euro zone[3], the UK and the US are shown in figure 2.1. Inflation rates are taken from the OECD database, except for Euro zone inflation, which was taken from Eurostat. Note that there are two outliers in the expectations time-series, namely in France (third quarter of 2000) and in Italy (second quarter of 1996), for which we controlled in our empirical analysis below by adding a dummy variable to the regression. The occurrence of these outliers is an unavoidable shortcoming of all conversion methods, when at a given point in time the assumption of normally distributed survey responses is violated.[4] Inflation expectations from the CESifo WES are 6-months-ahead inflation expectations which are queried every three months in the first two weeks of January, April, July and October. In Henzel and Wollmershäuser (2005) we showed that the information set that is available to the survey respondents at the time they fill in the questionnaire is the past quarter (that is the first quarter for the questionnaires returned at the beginning of April, the second quarter for the questionnaires returned at the beginning of July, and so on). Thus, the April survey produces inflation expectations $\bar{F}_t \pi_{t+2}$, where t refers to the first quarter and $t+2$ to the third quarter. As in a quarterly Phillips curve model, the required expectation's horizon should be a quarter of a year, it would be more convenient to use the CESifo WES 6-months-ahead inflation expectations together with semiannual data. In order to see whether the frequency of the data

[2]If, for example, perceived inflation is 1% (10%), an expected increase of the inflation rate of 0.46 (1.64) percentage points is needed to make the survey respondents mark UP in the questionnaire. By contrast, a decrease of the inflation rate of 0.29 (1.61) percentage points must be expected to make them mark DOWN.

[3]Euro zone inflation expectations have been calculated as a weighted sum of the responses for the individual member countries. The weights are the country weights used by Eurostat to calculate the Harmonized Index of Consumer Prices for the Euro zone. See Henzel and Wollmershäuser (2005) for further details.

[4]In the case of France, for example, in the October 2000 survey 13 out of 21 respondents indicated UP and 7 indicated DOWN. The problem was that only 1 respondent expected inflation to remain the same, which is a clear violation of the normality assumption.

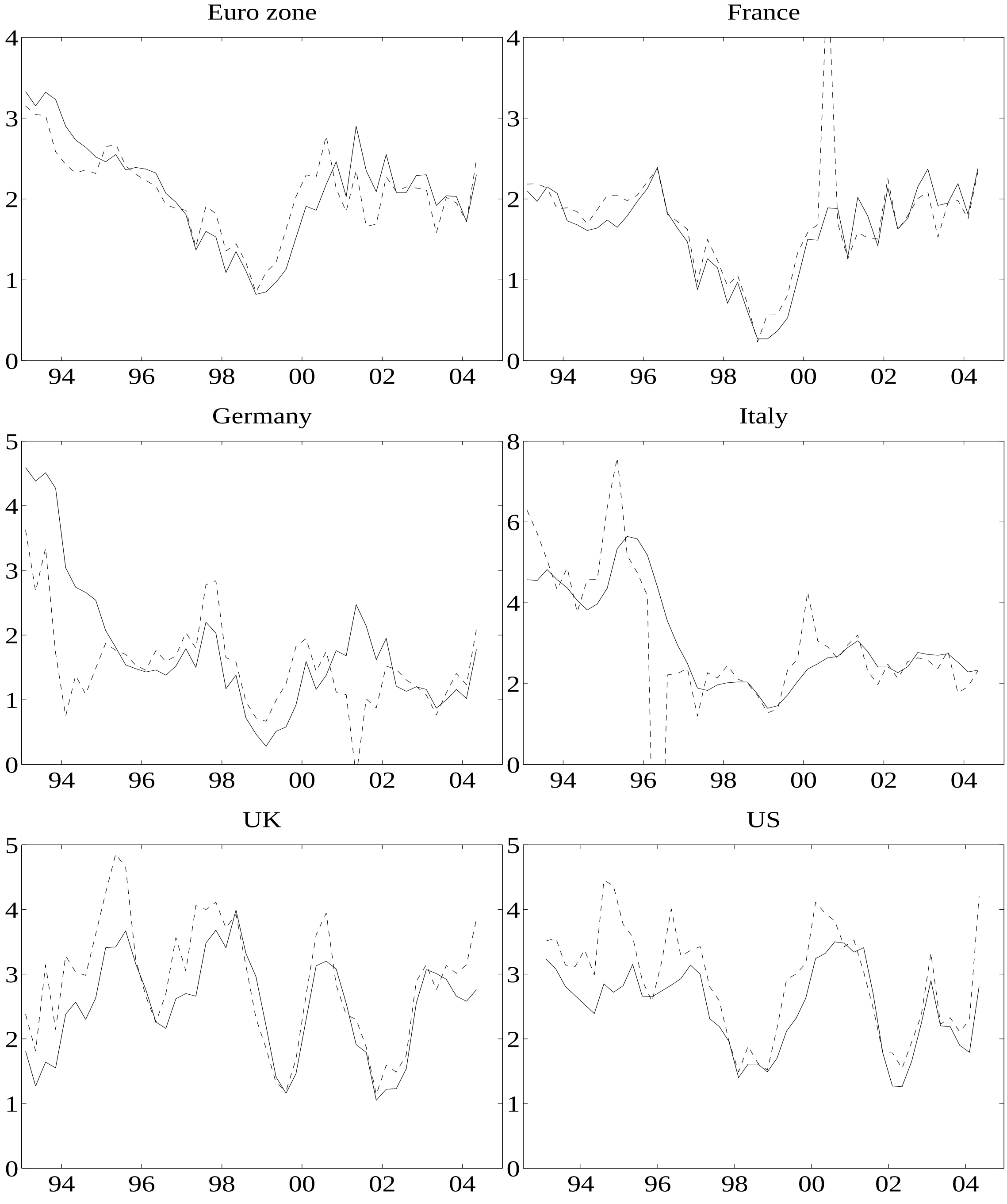


Figure 2.1: Actual inflation (π_t, solid line) and expected inflation ($\bar{F}_t \pi_{t+2}$, dashed line), in percent

matters for the empirical results, we ran regressions using both, quarterly and semi-annual data. As the estimated coefficients were almost identical, we decided to present only the results of the regressions that were obtained using quarterly data. By using the 6-months-ahead inflation expectations as proxies for 3-month expectations we implicitly assumed that forecaster's expectations are the same for each of the two upcoming 3-month periods (see also Roberts, 1997, on this point).

As most of the countries considered in this paper belonged to the European Monetary System, the data starts in the first quarter of 1993 in order to exclude the crisis which took place in September 1992. Compared to most other empirical Phillips curve studies this rather short estimation period is a novelty (see table 2.13 in appendix 2.F for a summary of other papers).

Using survey data for inflation expectations – instead of imposing rational expectations – when estimating a Phillips curve relationship should only produce different results, if survey expectations are not being formed rationally. The reason why we are questioning the rationality of survey expectations is due to the mixed evidence reported in the literature. Most papers that have examined survey measures of inflation expectations have concluded that these expectations are not rational in the sense of Muth (1961) (see for example Roberts, 1997, and the papers cited there).

A necessary condition for rational expectations is the unbiasedness of expectations. In order to find out whether CESifo WES expectations are unbiased predictors of future inflation, we regressed the forecast error (defined as $\pi_t - \bar{F}_{t-2}\pi_t$) on a constant c and tested whether it is significantly different from zero. Table 2.1 reveals that in the Euro zone, France and Germany, inflation expectations were unbiased during the period 1993:1 to 2004:2. By contrast, in the UK, the US and Italy expectations do not fulfill the necessary condition for rationality.[5] From the negative sign of the constant, we can conclude that expectations were biased upwards throughout the period of disinflation in the beginning of the 1990s.

A further necessary condition for rational expectations is the informational efficiency of expectations which requires the forecast error to be orthogonal to the information commonly available at time $t-2$. A first indication for the inefficiency of expectations is given by the p-values of the serial correlation LM-test in table 2.1, which indicate that – except for Italy and France – the residuals are not free of autocorrelation.[6] Autocorrelation in the forecast error implies that a shock to the

[5] Using the Livingston Survey of Professional Forecasters (which queries quantitative inflation expectations) Adam and Padula (2003) also find that expectations in the US were biased during the nineties.

[6] As the forecast horizon does not correspond to the frequency of the survey, shocks to the inflation rate can not be taken into account until the second period after the forecast and the same error may be repeated again. Thus, autocorrelation of order one in the error constitutes no

	Euro zone	France	Germany	Italy	UK	US
c	0.05	−0.04	0.13	−0.27	−0.37	−0.43
	[0.46]	[0.66]	[0.49]	[0.03]	[0.02]	[0.00]
LM(2)	0.82	0.77	0.00	0.96	0.00	0.00
LM(4)	0.00	0.33	0.00	0.98	0.00	0.01

Notes: We set a dummy variable to control for the outliers in France (2000:2) and Italy (1996:1) which are due to the conversion of inflation expectations from qualitative into quantitative data. The p-values, which have been calculated using Newey-West standard errors to correct for overlapping forecast errors, are reported in brackets. The last two rows report p-values for an LM test for the null hypothesis of no autocorrelation up to the second and fourth lag. Sample period: 1993:1 - 2004:2.

Table 2.1: Unbiasedness of expectations

inflation rate or to some other economic variable was not taken into account when the inflation forecast was made and that the same mistake was repeated in subsequent periods. In other words, efficiency of expectations requires that the forecast could not have been improved by adding additional information. In order to test for this, the forecast error is regressed on a number of exogenous variables that are known at time $t-2$ and that are possibly relevant when forecasting inflation.[7] Table 2.2 reports p-values related to χ^2-statistics of a Wald test of the null hypothesis that the coefficients on the aforementioned lags of these regressors are jointly equal to zero. In the Euro zone, France, Germany, Italy and the UK lagged values of the forecast error can explain the movement of the forecast error at the five percent level, which is a hint that survey respondents seem to be sluggish when correcting their expectations after having recognized the last forecast error. Also past inflation rates are of explanatory use in all countries. This means that respondents underestimate the inertia of the inflation rate. In none of the countries except France the output gap has a significant influence, indicating that the respondents seem to take it into account when forming their expectations. By contrast, real unit labor costs seem

irrationality.

[7]Our proceeding basically follows Roberts (1997) who introduced as potentially omitted variables the output gap as a measure of overall economic activity (see section 2.3.2 for a definition), the inflation rate to capture the persistence of inflation, and the three-month interest rate as an indicator for the stance of monetary policy. Since unit root tests indicated that the interest rates are non-stationary, we used first differences. In addition to that, we included real unit labor costs (see section 2.3.2 for a definition) and lagged terms of the forecast error. The explanatory power of each group of variables (which comprises four lags of the variable under consideration) was tested separately. The forecast error, real unit labor cost and the output gap enter the regression only from $t-3$ on, for reasons of overlapping forecast errors and because we assume a publication lag of one quarter.

to be omitted in the Euro zone, France, and the US. The three-month interest rate helps explain the forecast error in Germany, Italy, UK and the US.[8] Thus, the efficiency tests show that the null hypothesis of orthogonality must be rejected for all countries. This evidence suggests that the polled experts did not make efficient use of all the information available at the time the expectations were formed. Taken together, the results of the unbiasedness tests and efficiency tests conducted here lead us to conclude that the survey expectations do not possess the properties implied by Muth's definition of rational expectations.

Country	Error lags 3 to 6	Inflation lags 2 to 5	Output gap lags 3 to 6	RULC lags 3 to 6	3M Rate lags 2 to 5
Euro zone	0.00	0.00	0.11	0.04	0.17
France	0.00	0.00	0.00	0.00	0.43
Germany	0.03	0.01	0.60	0.07	0.00
Italy	0.00	0.00	0.16	0.08	0.02
UK	0.04	0.00	0.35	0.22	0.00
US	0.86	0.00	0.21	0.03	0.04

Notes: The table shows p-values for a heteroscedasticity and autocorrelation consistent (HAC) Wald-test on joint significance of each group of lagged variables (Error = forecast error, RULC = real unit labor cost, 3M Rate = three-month nominal interest rate). The dummy variables are set as described in the notes to table 2.1. Sample period: 1993:1 - 2004:2.

Table 2.2: Efficiency tests

2.3.2 Measures for Real Marginal Costs

There has been an extensive discussion in the literature about the correct proxy for real marginal costs (see for example Galí and Gertler, 1999, Galí, Gertler, and López-Salido, 2001 and Sbordone, 2005). There are basically two candidates that are considered: real unit labor costs and the output gap. The hypothesis that real unit labor costs is a good proxy for real marginal costs can be justified by the assumption that the production technology is Cobb-Douglas and that capital is constant over time. Real marginal costs are then defined as the ratio of real wages to the marginal product of labor

$$MC_t = \frac{1}{\alpha}\frac{W_t N_t}{P_t Y_t} \tag{2.13}$$

[8]Roberts (1997) and the studies cited there also find no support of the efficiency hypothesis for the US. Adam and Padula (2003) come to the same conclusion. For the Euro zone Forsells and Kenny (2004) who investigated qualitative inflation expectations from the European Commission's Consumer Survey also find that expectation were not efficient during the nineties.

where α is the labor elasticity of production, W_t the nominal wage rate, N_t employment, P_t the price level, and Y_t aggregate output. The second term on the right-hand-side is typically referred to as the labor income share or real unit labor costs. Log-linearizing equation (2.13) around the steady state gives

$$mc_t = w_t + n_t - p_t - y_t \tag{2.14}$$

where lower case-letters denote the percentage deviation of a variable around its steady state. Thus, under the assumption that α is constant over time, equation (2.14) shows that real marginal costs and real unit labor costs move in a one-to-one relation around their steady state.

While real unit labor costs are a direct measure of a firm's real marginal costs, it can be shown that under certain conditions the output gap is a close proxy. We will not go into the details of the derivation of this relationship because it has been well documented in standard textbooks on monetary economics (see for example Walsh, 2003, chapter 5.4). The idea is that after combining the households' labor supply decision (real wage equals the marginal rate of substitution between consumption and labor) with the firms' price-setting condition (price equals a mark-up over nominal marginal costs), an expression for the output level under both flexible and rigid prices can be derived. Under the assumption that labor market frictions exist but do not vary over time, real marginal costs are then a linear function of the output gap x_t

$$mc_t = (\sigma + \eta)(y_t - y_t^{flex}) = (\sigma + \eta)x_t \tag{2.15}$$

where $1/\sigma$ is the intertemporal elasticity of substitution in consumption, η the elasticity of marginal disutility with respect to labor supply, and y_t^{flex} the log of the level of output that would prevail if prices were perfectly flexible (i.e. $\theta = 0$). The HNKPC then becomes

$$\pi_t = \gamma_f \bar{F}_t[\pi_{t+1}] + \gamma_b \pi_{t-1} + \lambda' x_t \tag{2.16}$$

where $\lambda' = \lambda(\sigma + \eta)$.

In our empirical analysis we consider both types of measures for real marginal costs. Specifically we use

- the deviation of the logarithm of CPI-deflated unit labor costs (of the total economy)[9] from a linear trend (over the period 1990:1-2004:3): $RULC_t$;

[9] Unit labor costs of the total economy are taken from the OECD database. Italian unit labor costs are only available for the business sector (which is defined as total economy minus public sector).

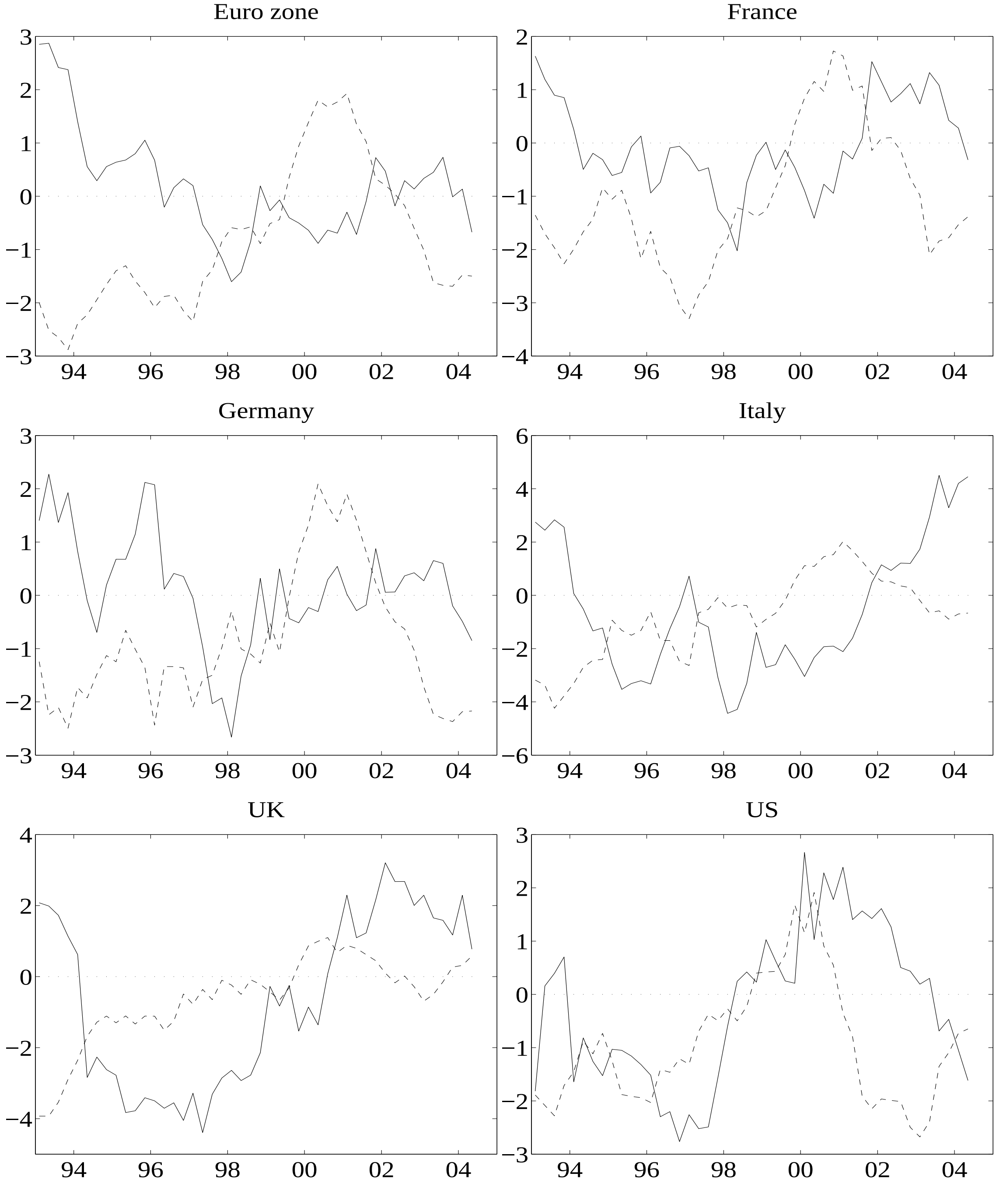


Figure 2.2: GAP_t (dashed line) and $RULC_t$ (continuous line), in percent

- and the OECD output gap (as published in the OECD Economic Outlook, Vol. 2004/2, No. 76):[10] GAP_t.

For each of the countries in our study, figure 2.2 shows both measures in a single graph.

2.4 Empirical Results

We begin by presenting estimates for the purely forward-looking New Keynesian Phillips curve which can be derived as a special case from the HNKPC by setting $\omega = 0$ (see table 2.3).

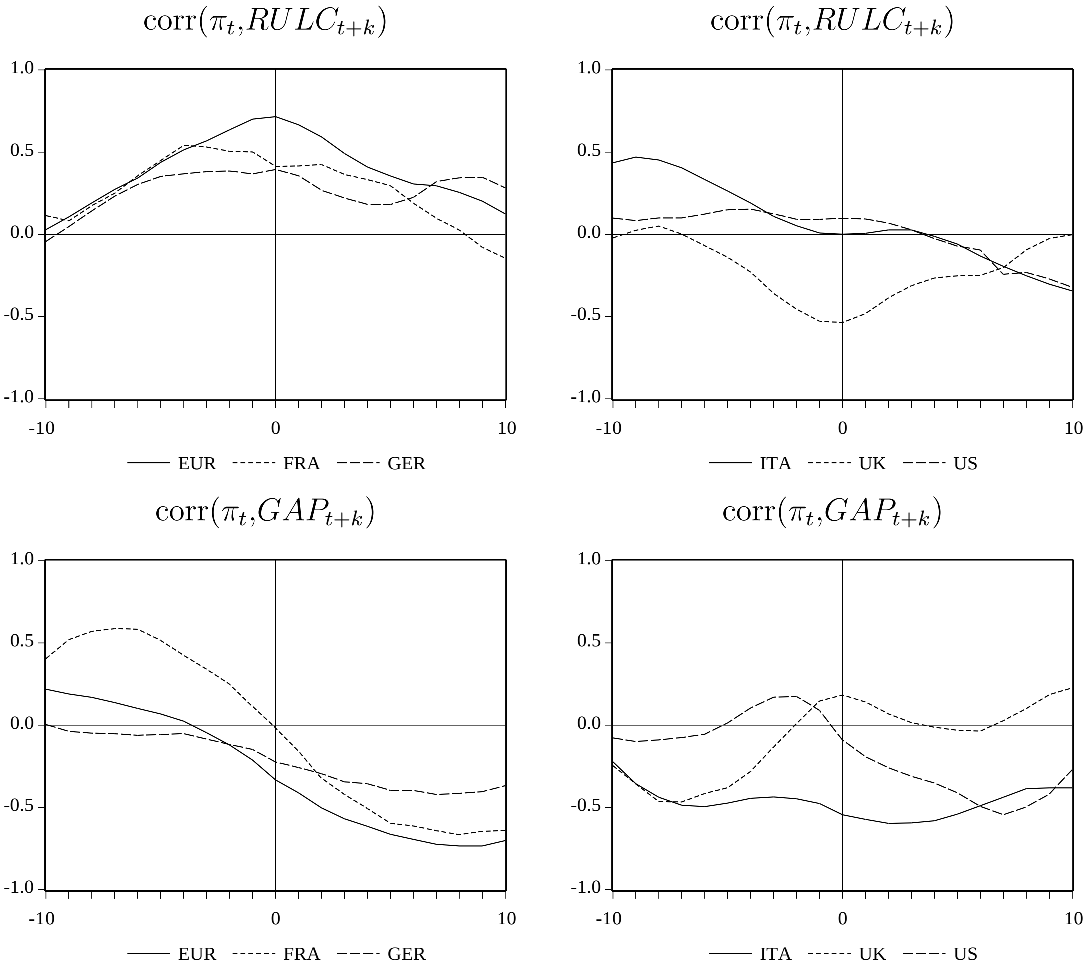
	unrestricted				restricted ($\beta = 0.99$)	
	RULC		GAP		RULC	GAP
	β	λ	β	λ'	λ	λ'
Euro zone	1.03	0.16	1.02	−0.08	0.19	−0.10
	[0.00]	[0.00]	[0.00]	[0.25]	[0.00]	[0.02]
France	0.97	0.08	1.00	0.03	0.07	0.03
	[0.00]	[0.03]	[0.00]	[0.24]	[0.05]	[0.19]
Germany	1.07	0.36	1.05	−0.09	0.37	−0.12
	[0.00]	[0.02]	[0.00]	[0.62]	[0.03]	[0.46]
Italy	0.91	0.02	0.89	−0.07	0.05	0.05
	[0.00]	[0.62]	[0.00]	[0.41]	[0.41]	[0.58]
UK	0.85	−0.00	0.87	0.11	0.07	0.22
	[0.00]	[0.99]	[0.00]	[0.01]	[0.19]	[0.00]
US	0.85	0.07	0.82	−0.06	0.13	0.14
	[0.00]	[0.16]	[0.00]	[0.13]	[0.13]	[0.04]

Notes: Numbers in brackets are p-values which were calculated using HAC Newey-West standard errors. For France and Italy we set a dummy variable in 2000:2 and 1996:1, respectively (see section 2.3.1). Sample period: 1993:1 - 2004:2.

Table 2.3: Estimation results for the forward-looking Phillips curve

[10] The OECD output gap is measured as the percentage difference between actual GDP in constant prices and potential GDP. The latter is estimated by the OECD using a production function approach, taking into account the capital stock, changes in labor supply, factor productivity and underlying non–accelerating wage rates of unemployment (see Giorno et al., 1995). An alternative measure of the output gap obtained from a Hodrick–Prescott–filtered GDP series gave results, which are less favorable for the Phillips curve. While in the majority of cases the signs of the slope coefficients were consistent with those obtained using the OECD output gap – in particular for the hybrid NKPC – almost all slope coefficients turned out to be insignificant.

In the purely forward-looking case the estimated parameter of inflation expectations is equal to the discount factor β. Irrespective of the model specification the β's are all statistically significant and in the neighborhood of one. Concerning the slope coefficient λ our results are to some extent in line with those obtained by Galí, Gertler, and López-Salido (2001) who used a rational expectations-GMM approach. While some of the estimated λ's of the RULC model (Euro zone, France and Germany) are positive and significant, the rest of the countries show no significant effect. The λ' of the output gap model is positive and significant only for the UK. For all other countries (except for France), it turns out to be negative, but the influence on inflation is insignificant. These results are perfectly in line with the cross correlations between inflation and RULC on the one hand, and inflation and the output gap on the other hand (see figure 2.3). For $k = 0$ (that is, contemporaneous correlation) λ is positive and significant only in those cases where correlations in figure 2.2 are positive as well. Table 2.3 also shows that the results improve when the estimation of the Phillips curve is restricted to $\beta = 0.99$. Except for the Euro zone and Germany the negative slope coefficients become positive, and for the US even significantly so.



Note: The correlation coefficient is depicted on the vertical axis, and k on the horizontal axis.

Figure 2.3: Cross correlograms

For most countries, however, the estimations of the purely forward-looking New Keynesian Phillips curve produce residuals with a high degree of autocorrelation, which indicates that some important explanatory variables are missing. We therefore turn to the estimation of the HNKPC which explicitly allows lagged inflation to have additional explanatory power for current inflation.[11] Table 2.4 reveals that in all of our estimations of the HNKPC the coefficients for both subjective inflation expectations and lagged inflation are positive and significant. For most countries the point estimates of γ_b turn out to be higher in the output gap model, whereas the γ_f's are somewhat lower. Looking at the individual countries, we can distinguish between three groups. In Germany and Italy the degree of backwardness is relatively high. Irrespective of the measure for marginal costs, γ_b exceeds γ_f. In France the opposite is true. The estimated γ_f's are higher than the γ_b's, implying that French firms are more forward-looking than their German or Italian competitors. In the Euro zone as a whole, the US and – to some extent – the UK optimizing firms and rule-of-thumb price setters are more or less balanced.

	RULC			GAP		
	γ_f	γ_b	λ	γ_f	γ_b	λ'
Euro zone	0.55	0.47	0.07	0.51	0.52	−0.00
	[0.00]	[0.00]	[0.04]	[0.00]	[0.00]	[0.94]
France	0.78	0.21	0.06	0.76	0.25	0.04
	[0.00]	[0.00]	[0.03]	[0.00]	[0.00]	[0.01]
Germany	0.29	0.74	0.08	0.28	0.78	0.04
	[0.00]	[0.00]	[0.27]	[0.00]	[0.00]	[0.30]
Italy	0.25	0.72	−0.01	0.26	0.74	0.05
	[0.00]	[0.00]	[0.47]	[0.00]	[0.00]	[0.00]
UK	0.49	0.43	−0.01	0.52	0.42	0.10
	[0.00]	[0.00]	[0.83]	[0.00]	[0.00]	[0.00]
US	0.46	0.46	0.04	0.44	0.48	0.01
	[0.00]	[0.00]	[0.07]	[0.00]	[0.00]	[0.70]

Notes: See table 2.3.

Table 2.4: Estimation results for the hybrid Phillips curve

The sign and significance of the measure for real marginal costs crucially de-

[11] Instead of introducing nominal rigidities in the price setting mechanism (i.e. rule-of-thumb price setters) autocorrelated residuals could also be explained by persistent shocks to the mark-up of firms over nominal marginal costs. Both modeling strategies would account for the empirical fact that inflation is inertial (see de Walque et al., 2006, for a recent paper dealing with this issue). While mark-up shocks seem to be the dominant source of inflation variability, their persistence alone fails to capture the empirical evidence from VAR studies that the response of inflation to shocks is gradual and hump-shaped (see Estrella and Fuhrer, 2002). According to the persistent mark-up theory all firms completely front load changes in prices in response to news about future profits, which leads to the typical jump behavior of aggregate inflation.

pend on the empirical specification of the HNKPC and differ from the results obtained from the estimation of the purely forward-looking Phillips curve. The most striking result is that the output gap becomes an important explanatory variable for inflation in France, Italy and – as in the forward-looking version – the UK, which is astonishing, given the low and mostly negative contemporaneous correlation between the output gap and inflation (see figure 2.3). From an econometric point of view, the significant output gap coefficients can be explained by the high correlation of the output gap with the unexplained part of a regression of inflation on lagged and expected inflation. When RULC are used as a measure for marginal costs, the results are more or less in line with those for the purely forward-looking Phillips curve. The λ's for the Euro zone and France remain positive and significant, whereas λ for Germany becomes insignificant. These results are roughly in line with the cross correlations we present in the upper part of figure 2.3.

	RULC		GAP	
	γ_f	λ	γ_f	λ'
Euro zone	0.52	0.08	0.49	-0.02
	[0.00]	[0.00]	[0.00]	[0.32]
France	0.78	0.05	0.76	0.03
	[0.00]	[0.05]	[0.00]	[0.04]
Germany	0.26	0.09	0.22	0.00
	[0.00]	[0.24]	[0.00]	[0.94]
Italy	0.25	-0.01	0.26	0.06
	[0.00]	[0.70]	[0.00]	[0.00]
UK	0.44	0.02	0.48	0.14
	[0.00]	[0.42]	[0.00]	[0.00]
US	0.35	0.05	0.36	0.08
	[0.00]	[0.20]	[0.00]	[0.01]

Notes: See table 2.3.

Table 2.5: Estimation results for the hybrid Phillips curve when $\gamma_f + \gamma_b = 1$

A necessary condition for the dynamic process to be stable is that the sum of γ_f and γ_b does not exceed one. To be sure that the process is not exploding, we also estimate a restricted version of the HNKPC where $\gamma_f + \gamma_b = 1$. Table 2.5 shows that the estimates of γ_f are still highly significant in every country. For the Euro zone, France, Germany and Italy the imposed restriction leaves the estimates of both, γ_f and λ/λ' more or less unchanged. By contrast, the restriction leads to a positive slope in the RULC model for the UK and a positive and significant slope in the GAP model for the US.

2.5 Discussion of the Results

The following results can be summarized from the previous section. First, the HNKPC performs better than the purely forward-looking Phillips curve. Not only the estimates for γ_f, but also those for γ_b are positive and significantly different from zero. Thus, there is a forward-looking and a backward-looking component of inflation dynamics, which exhibits considerable variation between countries. Second, it is unclear whether real unit labor costs or the output gap should be used as proxy for real marginal costs. While for some countries (the Euro zone, France, Germany) the first seems to be the driving variable of inflation, for other countries (Italy, the UK, the US) the output gap is the appropriate measure. This shows that both are only an imperfect proxy for the typically unobserved real marginal costs. Third, theory-based restrictions ($\beta = 0.99$ and $\gamma_f + \gamma_b = 1$) lead to an improvement of the empirical results. This is especially the case for the US and, to some extent, the UK where the unrestricted estimates of β and $\gamma_f + \gamma_b$ are significantly lower than 0.99 and 1.

In the following, we compare our results with those obtained in other empirical studies. The direct use of measures for inflation expectations, which naturally avoids any assumptions on the expectations formation process is much less popular than the rational expectations approach. We only found five studies using either survey data or OECD forecasts for expectations, which are summarized in table 2.12 in appendix 2.F. The great majority of empirical work on the New Keynesian Phillips curve applies the rational expectations approach. Table 2.13 in appendix 2.F presents some of the most recent papers.

An interesting result of our paper is that, except for France, the degree of forward-looking behavior is found to be lower when using survey data instead of imposing rational expectations. And this finding is qualitatively confirmed by the other survey data studies. In table 2.6 we calculated averages of the estimates of γ_f that are presented in tables 2.12 and 2.13 in appendix 2.F. Germany is a very striking example. While studies using the rational expectations approach find an average coefficient for γ_f of 0.67, our estimates are much lower, with an average value of 0.26. Reckwerth (1997) who uses another source for German inflation expectations also finds estimates for γ_f which are smaller than under the rational expectations approach. The results for the US point into the same direction. While the average value for γ_f under rational expectations is 0.61, our regressions returned an average value of 0.40. Again, this tendency of a lower degree of forward-looking behavior is confirmed by other studies using survey data.

There are two possible explanations for the gap between the estimates for γ_f. First, we provide evidence for a more recent period as our sample starts at

	Euro zone	France	Ger -many	Italy	UK	US
Ifo WES	0.52	0.77	0.26	0.26	0.48	0.40
other surveys	0.49	-	0.43	-	-	0.43
RE approach	0.62	0.65	0.67	0.52	0.70	0.61

Notes: As some of the papers cited in table 2.13 of appendix F only report standard errors for the deep parameters of the Phillips curve (β, θ, ω), but not for γ_f, we resorted to a non-parametric statistical significance test, the Wilcoxon rank-sum test, for testing whether our γ_f's are significantly lower than those in the RE (= rational expectations) literature. The results, which are available from the authors upon request, show that the null hypothesis that the median of the RE estimates is lower than our estimates can be rejected at the 5% level for all countries except France.

Table 2.6: Summary of estimates for γ_f

the beginning of the 1990s whereas most of the other studies begin in the 1960s or 1970s. Most of the countries in our sample, however, underwent one or even more significant changes in their monetary policy strategy. As the monetary policy regime that is in force plays a crucial role for the estimated behavioral parameters, it is likely that these models suffer from instabilities that cannot be accounted for by GMM techniques. Unfortunately, stability of the results is rarely discussed in these papers. Note, however, that a shorter sample size cannot explain why γ_f in these studies tends to be systematically higher than in our study. A further argument against this explanation are the results of other studies listed in table 2.12 using US survey data. While their samples range from the 1960s to 1999 or later, which roughly corresponds to to the time span covered by most rational expectations studies, the average point estimates for γ_f are close to ours.

Second, the above findings are consistent with some recent papers questioning the appropriateness of the rational expectations approach. The Phillips curve model is typically estimated by replacing expectations with actual realizations and by deriving orthogonality conditions that may be used to estimate the parameters of the model with GMM. These moment conditions are derived on the assumption that expectations are rational, i.e. that the expectation-induced 'errors in variables' must be orthogonal to the information set available to the agents at the time the expectations are formed. Rudd and Whelan (2005), however, argue that GMM estimates may overstate the degree of forward-looking behavior if the instrument set includes variables that directly cause inflation but are omitted from the hybrid model specification. Thus, the error term of the estimation equation is no longer a pure rational expectations error. As a consequence estimates for γ_f will be biased upwards as long as both, π_{t+1} and its instruments, are correlated with the omitted variable. According to Mavroeidis (2005) the mis-specification of the Phillips curve model can alternatively be interpreted as a failure of rationality. Irrespective of the source for the violation of the orthogonality conditions, he shows that the model is weakly identified, which introduces a bias in the GMM estimation in favor of a hybrid specification with apparently dominant forward-looking behavior.

The use of survey data avoids this problem as there is no need to specify the expectation formation process and expectations can be taken as exogenous. Section 2.2.2 shows in detail that a Phillips curve relationship can be derived for a wide range of expectations formation processes. The only assumption that is made, is that the law of iterated expectations holds. Rational expectations can be considered as a special case. According to this theoretical model, which serves as a basis for the econometric part of the paper, agents are considered to be forward-looking if they do not simply extrapolate past inflation rates into the future, but use additional information when they reset prices. This does not necessarily mean that expectations have to be rational. They can be anything else except purely backward-looking (i.e. $E_t\pi_{t+1} = \pi_{t-1}$), which would lead to perfect multicollinearity, and hence to problems in the OLS estimation. In section 2.3.1 we showed that inflation expectations from the CESifo WES are far from being rational. In addition, we ran some simple regressions of expected inflation on past inflation, which indicate that survey expectations contain information way beyond past inflation rates.[12] In section 2.4 the behavior of forward-looking agents is then approximated by the survey expectations of the CESifo WES. Under the assumption that these are an accurate measure of ‘true’ expectations, we estimated the share of forward-looking agents using the subjective expectations HNKPC and showed that there is a role for forward-looking behavior. Given the theoretical set-up this result holds irrespective of the information content of the survey expectations.

The advantage of the survey data approach is that its results are more reliable than those obtained from the GMM approach. On the one hand, there is no need for instruments. Appendix 2.D shows at length that the estimations of the HNKPC do not suffer from a violation of the OLS assumptions. Ruling out weak instruments we re–estimate the Phillips Curve by TSLS and find that, in particular, endogeneity of the regressors can be rejected. On the other hand, the OLS estimation has the advantage that the results are robust despite the relatively short sample size. A CUSUM of squares test, which is presented in appendix 2.E, yields satisfactory results for the stability of the estimated parameters. It is well known that the small sample properties of GMM estimations are very poor, meaning that estimators are often found to be biased, widely dispersed and sensitive to the normalization of the orthogonality conditions as well as to the choice of the instruments (see for

[12]Specifically, we find that π_{t-1} only explains between 16% (for Germany) and 56% (for France) of the variation of expectations $\bar{F}_t\pi_{t+2}$. These results are not shown in the paper, but available from the authors upon request

example Fuhrer, Moore, and Schuh, 1995, on this issue). This is also the reason why we did not apply the rational expectations GMM approach to a shorter sample (starting for example in 1993) in order to find out whether the identified gap between the estimates for γ_f are due to the sample size or due to the proxy for inflation expectations and the related estimation methodology.

2.6 Conclusions

In this paper we provided evidence for the fit of the hybrid New Keynesian Phillips curve for selected Euro zone countries, the US and the UK. On a theoretical level, we derived the Phillips curve under the assumption that a fraction of firms has subjective expectations that may be non-rational, while the remaining fraction uses a simple rule-of-thumb. On an empirical level, we estimated the Phillips curve by using direct measures of inflation expectations from the CESifo World Economic Survey, instead of imposing rational expectations and estimating the Phillips curve by GMM.

Our main findings are as follows: First, the use of survey data gives empirical results, which are more reliable than those obtained from the rational expectations GMM approach. With survey data there is no need for instruments and this provides consistently and efficiently estimated parameters even when there is only a limited sample available to the researcher. We show that OLS parameters are stable over time and endogeneity of regressors can be rejected. Second, the purely forward-looking Phillips curve can be rejected in favor of the hybrid New Keynesian Phillips curve as the estimated coefficients on past inflation are significantly different from zero. Third, the estimated coefficients on past inflation are higher when using survey expectations than when using the rational expectations GMM approach. This result is consistent with some recent papers, which argue that GMM estimates for the degree of forward-looking behavior are biased upwards if the Phillips curve model is mis-specified. Fourth, it remains unclear whether real unit labor costs or a measure of the output gap should be used as a proxy for real marginal costs. For some countries (the Euro zone, France, Germany) the first is better, while for other countries (Italy, the UK, the US) the latter is superior. This shows that both are only an imperfect proxy for real marginal costs. Fifth, theory-based restrictions lead to an improvement of the empirical results.

One explanation for our findings is that non-rationalities which are incorporated in survey expectations may matter for the price-setting process of firms. If we are correct in using a survey among economic experts for approximating firms' expectations, such an explanation would have an important impact on the policy conclusions that are typically drawn on the basis of general equilibrium models where agents are assumed to form expectations rationally. Some first attempts to model deviations from perfectly rational expectations have been developed by Mankiw and Reis (2002). In their sticky-information model they impose a constraint on the information that people use when forming expectations. They assume that in each period there is a fixed probability that a person updates his information set; otherwise he continues to set prices on outdated information. In Ball, Mankiw, and Reis (2005) they provide a normative monetary policy analysis that accounts for these deviations from rationality. And their central conclusion is that under such a setting the central bank should target the price level rather than the inflation rate. Thus, in future work it would be interesting to investigate in more detail how the private sector actually forms inflation expectations.

Appendix

2.A Derivation of Equation (2.9)

Equation (2.9) can be derived by subtracting equation (2.7) from equation (2.6)

$$F_t^i[p_{t+1}^f] - p_t^f = F_t^i[\pi_{t+1}^f] = \frac{1}{I} F_t^i[\sum_{h=1}^{I}(p_{t+1}^{f,h} - p_t^{f,h})]$$

and by replacing $p_{t+1}^{f,h}$ and $p_t^{f,h}$ with the first expression of equation (2.5):

$$\begin{aligned} F_t^i[\pi_{t+1}^f] &= \frac{1-\beta\theta}{I} F_t^i[\sum_{h=1}^{I}\{F_{t+1}^h[\sum_{k=0}^{\infty}(\beta\theta)^k(mc_{t+k+1} + p_{t+k+1})] - \\ &\quad -F_t^h[\sum_{k=0}^{\infty}(\beta\theta)^k(mc_{t+k} + p_{t+k})]\}]. \end{aligned}$$

Applying the law of iterated expectations (equation (2.8)) this expression can be simplified to

$$F_t^i[\pi_{t+1}^f] = \frac{1-\beta\theta}{I} F_t^i[\sum_{h=1}^{I}\{(1-\beta\theta)F_t^h[\sum_{k=0}^{\infty}(\beta\theta)^k(mc_{t+k+1} + p_{t+k+1})] - (mc_t + p_t)\}].$$

Replacing $(1-\beta\theta)F_t^h[\sum_{k=0}^{\infty}(\beta\theta)^k(mc_{t+k+1} + p_{t+k+1})]$ with $p_{t+1}^{f,h}$ (equation (2.5)) and using equation (2.6) finally gives equation (2.9):

$$F_t^i[\pi_{t+1}^f] = (1-\beta\theta)(F_t^i[p_{t+1}^f] - mc_t - p_t).$$

2.B Derivation of Equation (2.10)

Equation (2.10) can be derived by aggregating equation (2.3) to

$$p_t^* = (1-\omega)p_t^f + \omega p_t^b,$$

solving the resulting expression for p_t^f and replacing p_t^b with equation (2.4):

$$p_t^f = \frac{p_t^* - \omega(p_{t-1}^* + p_{t-1} - p_{t-2})}{1-\omega}.$$

Next, solve equation (2.2) for p_t^* and replace it in the preceding expression. After a little algebra, equation (2.10) is obtained:

$$p_t^f = \frac{p_t + (\theta\omega - 2\omega - \theta)p_{t-1} + \omega p_{t-2}}{(1-\theta)(1-\omega)}.$$

2.C Derivation of Equation (2.12)

Inserting equation (2.11) on the right-hand-side of equation (2.9) gives

$$F_t^i[\pi_{t+1}^f] = (1-\beta\theta)\left(\frac{F_t^i[\pi_{t+1}] - \omega\pi_t}{(1-\theta)(1-\beta)} - mc_t\right).$$

Forming average subjective expectations, $\bar{F}_t[\cdot] = (1/I)\sum_{i=1}^{I} F_t^i[\cdot]$, yields

$$\bar{F}_t[\pi_{t+1}^f] = (1-\beta\theta)\left(\frac{\bar{F}_t[\pi_{t+1}] - \omega\pi_t}{(1-\theta)(1-\beta)} - mc_t\right).$$

An alternative expression for $F_t^i[\pi_{t+1}^f]$ can de derived by subtracting equation (2.10) from equation (2.11):

$$F_t^i[\pi_{t+1}^f] = \frac{F_t^i[\pi_{t+1}] + (\theta\omega - 2\omega - \theta)\pi_t + \omega\pi_{t-1}}{(1-\theta)(1-\omega)}.$$

Forming average subjective expectations yields

$$\bar{F}_t[\pi_{t+1}^f] = \frac{\bar{F}_t[\pi_{t+1}] + (\theta\omega - 2\omega - \theta)\pi_t + \omega\pi_{t-1}}{(1-\theta)(1-\omega)}.$$

Equating both expressions for $\bar{F}_t[\pi_{t+1}^f]$ and solving for π_t finally results in equation (2.12):

$$\pi_t = \gamma_f \bar{F}_t[\pi_{t+1}] + \gamma_b\pi_{t-1} + \lambda mc_t,$$

where

$$\gamma_f = \frac{\beta\theta}{\theta + \omega[1-\theta(1-\beta)]},$$

$$\gamma_b = \frac{\omega}{\theta + \omega[1-\theta(1-\beta)]},$$

$$\lambda = \frac{(1-\omega)(1-\theta)(1-\beta\theta)}{\theta + \omega[1-\theta(1-\beta)]}.$$

2.D Endogeneity of the Regressors

This appendix investigates whether the estimates presented in tables 2.3 to 2.5 suffer from endogeneity of the regressors. On the one hand, the expectational variable $\bar{F}_t\pi_{t+2}$ may be caused by the current inflation rate – a problem which is often referred to as simultaneity. On the other hand, the OLS estimations of the HNKPC may suffer from correlation between the lagged endogenous variable π_{t-1} and the error term of the regression.

The endogeneity problem of the OLS regressions is addressed by re-estimating the two Phillips curve models using instrumental variables methods. Specifically, we run a two-stage least squares (TSLS) regression where we instrument for $\bar{F}_t\pi_{t+2}$ and – in the case of the HNKPC – for π_{t-1}. As instruments for $\bar{F}_t\pi_{t+2}$ we consider a constant and up to six lags of inflation expectations, real unit labor costs and the output gap, which are the driving variables of the inflation process. Six lags should be sufficient to account for the dynamics in the economy. In the case of the HNKPC we are likely to face a difficulty due to the fact that both endogenous regressors (expected and lagged inflation) often follow a very similar pattern. By simply using a constant and lagged values of inflation expectations, real unit labor costs and the output gap, the instruments may be weak in the sense of Stock and Yogo (2005) and thus lead to a bias in the TSLS estimates. Thus, we need to choose a set of instrumental variables that allows us to identify the effect of both, inflation expectations and lagged inflation, on current inflation.

In order to test whether the instrumental variables estimations suffer from weak instruments, we calculate the Cragg-Donald statistic for all possible combinations of instruments (i.e. 2^{19} combinations).[13] In the case of Italy we additionally have to include lags of the spread between the short-term and the long-term interest rate as a possible instrumental variable to rule out weak instruments. For each number of instruments we then record the combination of instruments that results in the maximum value of the Cragg-Donald statistic. Finally, we choose the instrument

[13] The test statistic was proposed by Cragg and Donald (1993) and Stock, Wright, and Yogo (2002). It can be seen as the matrix analog of the first-stage F-statistic in a multivariate setting.

set, which contains the largest number of instruments and, at the same time, gives rise to a bias of the IV regression compared to OLS which is significantly less than 10% according to table 1 in Stock and Yogo (2005). This algorithm chooses the instrument sets described in table 2.7. Table 2.8 presents the related Cragg-Donald statistics, which shows that in all cases the null hypothesis of weak instruments can be rejected at the 5% significance level.

	c	$\bar{F}_{t-k}\pi_{t-k+2}$						GAP_{t-k}						$RULC_{t-k}$					
k		1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
								RULC											
Euro zone	*			*	*			*		*				*				*	
France					*		*												
Germany	*	*				*			*		*	*					*	*	
Italy$^\diamondsuit$				*							*				*			*	
UK	*		*	*														*	*
US					*	*	*			*									
								GAP											
Euro zone	*			*	*			*		*				*				*	
France					*		*												
Germany	*		*		*	*					*	*					*	*	*
Italy$^\diamondsuit$		*										*			*				
UK	*			*								*				*			
US					*	*					*						*		

Notes: A "$*$" indicates that the respective lag (or constant) was included in the instrument set.
$\diamondsuit$: For the estimation of the HNKPC of Italy lags 1 and 4 of the interest rate spread enter the estimation as instruments in addition to the ones marked with a "$*$".

Table 2.7: Instrument sets for the HNKPC

In a next step we present the results of the TSLS regressions in tables 2.9 to 2.11. Provided that the instruments are valid, it is possible to test whether one of the regressors is endogenous. Hausman (1978) proposed a test, which compares the estimates of the OLS regression with those of the TSLS regression. If they are not systematically different, one should rely on the results of the OLS regression and conclude that there is no problem of endogenous regressors.[14] In tables 2.9 to 2.11 the columns, which are labeled with H show the p-values for the null hypothesis that the OLS estimates are consistent and efficient.

[14] Note that the distribution of the test statistic is only known for variance-covariance matrices of estimators that have not been adjusted for heteroscedasticity and autocorrelation. Therefore, we use the unadjusted variance-covariance matrices to calculate the test statistic. However, the p-values for the significance of the estimated parameters were calculated using Newey-West adjustment of the standard errors.

	Euro zone	France	Ger -many	Italy	UK	US
	RULC					
forward-looking	145.0‡ [19]	68.02‡ [19]	32.90‡ [19]	32.90‡ [19]	67.01‡ [19]	77.07‡ [19]
hybrid	10.03† [7]	8.82† [2]	10.68† [8]	9.13† [6]	10.32† [5]	7.82† [4]
	GAP					
forward-looking	107.7‡ [19]	34.16‡ [19]	19.91† [19]	23.30‡ [19]	53.39‡ [19]	50.37‡ [19]
hybrid	10.63† [7]	8.35† [2]	10.89† [9]	9.08† [5]	9.25† [4]	8.40† [4]

Notes: The number of instruments used in each regression is given in brackets. Critical values for the 5% significance level are provided by Stock and Yogo (2005), table 1, with ‡ (†) denoting a desired maximal bias of the IV estimator relative to OLS of 5% (10%). Sample period: 1993:1 - 2004:2.

Table 2.8: Weak instrument test (Cragg-Donald statistic)

	RULC			GAP		
	β	λ	H	β	λ'	H
Euro zone	1.03 [0.00]	0.16 [0.01]	0.09	1.03 [0.00]	-0.07 [0.25]	0.12
France	0.98 [0.00]	0.08 [0.03]	0.70	1.00 [0.00]	0.03 [0.17]	0.67
Germany	1.10 [0.00]	0.35 [0.01]	0.11	1.10 [0.00]	-0.05 [0.77]	0.14
Italy	0.94 [0.00]	0.03 [0.52]	0.01	0.92 [0.00]	-0.03 [0.72]	0.01
UK	0.86 [0.00]	0.00 [0.93]	0.11	0.88 [0.00]	0.12 [0.01]	0.06
US	0.85 [0.00]	0.07 [0.13]	0.07	0.83 [0.00]	-0.05 [0.16]	0.13

Notes: See table 2.3. The column labeled with H shows the p-values for the null hypothesis that the OLS estimates are consistent and efficient.

Table 2.9: IV (TSLS) estimation results for the forward-looking Phillips curve

The following results can be summarized. A comparison of table 2.9 with table 2.3 shows that the estimated coefficients only change slightly. While, as a general rule, the estimates become somewhat higher, the significance of the coefficients remains unaffected. Nevertheless, the null hypothesis of the Hausman test is rejected at the 5% level for Italy (irrespective of the proxy for real marginal costs) and at the 10% level for the RULC model of the Euro zone and the US and the GAP model of the UK. This may be due to the fact that we apply the test to OLS and TSLS estimates whose variance-covariance matrix is not adjusted for heteroscedasticity and autocorrelation. As the latter is only a problem when estimating the purely forward-looking Phillips curve, the results of the Hausman test have to be interpreted with caution.

	RULC				GAP			
	γ_f	γ_b	λ	H	γ_f	γ_b	λ'	H
Euro Zone	0.55	0.47	0.07	0.62	0.48	0.55	0.00	0.63
	[0.00]	[0.00]	[0.04]		[0.00]	[0.00]	[0.91]	
France	0.62	0.36	0.05	0.54	0.71	0.30	0.04	0.93
	[0.00]	[0.00]	[0.08]		[0.00]	[0.00]	[0.02]	
Germany	0.28	0.76	0.08	0.58	0.26	0.81	0.05	0.19
	[0.00]	[0.00]	[0.28]		[0.00]	[0.00]	[0.24]	
Italy	0.29	0.69	−0.01	0.35	0.29	0.72	0.06	0.20
	[0.00]	[0.00]	[0.65]		[0.00]	[0.00]	[0.02]	
UK	0.47	0.45	−0.01	0.55	0.54	0.39	0.10	0.89
	[0.00]	[0.00]	[0.74]		[0.00]	[0.00]	[0.00]	
US	0.50	0.41	0.04	0.48	0.41	0.51	0.01	0.29
	[0.00]	[0.00]	[0.06]		[0.00]	[0.00]	[0.87]	

Notes: See table 2.3. The column labeled with H shows the p-values for the null hypothesis that the OLS estimates are consistent and efficient.

Table 2.10: IV (TSLS) estimation results for the hybrid Phillips curve

In the case of the HNKPC (tables 2.10 and 2.11) the results of the Hausman test indicate unambiguously that the OLS estimates are reliable. A comparison of table 2.10 with table 2.4 shows that the TSLS estimates are very close to those obtained from the OLS regression. As before the significance of the coefficients remains unaffected. A similar conclusion can be drawn for the restricted HNKPC (tables 2.11 and 2.5). As a general rule, the estimated coefficients (and hence the degree of forward-looking behavior) are somewhat lower than in the OLS regression. To sum up, the TSLS estimates and the results of the Hausman test indicate that the OLS procedure is superior to TSLS and yields reliable estimates for the Phillips curve.

	RULC			GAP		
	γ_f	λ	H	γ_f	λ'	H
Euro zone	0.51	0.07	0.94	0.45	-0.02	0.73
	[0.00]	[0.00]		[0.00]	[0.42]	
France	0.60	0.04	0.45	0.60	0.03	0.34
	[0.00]	[0.11]		[0.00]	[0.04]	
Germany	0.23	0.08	0.69	0.18	0.01	0.45
	[0.00]	[0.26]		[0.00]	[0.81]	
Italy	0.19	-0.01	0.68	0.30	0.06	0.70
	[0.09]	[0.55]		[0.00]	[0.00]	
UK	0.37	0.01	0.31	0.45	0.13	0.81
	[0.00]	[0.59]		[0.00]	[0.00]	
US	0.29	0.04	0.42	0.27	0.07	0.11
	[0.00]	[0.30]		[0.00]	[0.05]	

Notes: See table 2.3. The column labeled with H shows the p-values for the null hypothesis that the OLS estimates are consistent and efficient.

Table 2.11: IV (TSLS) estimation results for the hybrid Phillips curve when $\gamma_f + \gamma_b = 1$

2.E Stability of the Estimates

This appendix investigates whether the estimated coefficients of the Phillips curve are stable over time. Figures 2.4 and 2.5 show the results of a CUSUM of squares test at the 1% level. For both, the purely forward-looking Phillips curves and the HNKPC, the cumulated sum of squares of recursive residuals lie within the significance lines for Germany, Italy, the UK and the US, implying that the regression relationship is constant over time, irrespective of the chosen measure for real marginal cost. For the other countries, the cumulated sum of squares of recursive residuals temporarily crosses the significance lines, but then stays inside the thresholds again. Note however that due to the high degree of auto-correlation in the residuals of the purely forward-looking Phillips curves, the CUSUM of squares test has to be interpreted with caution as the distribution of the test statistic is not known.

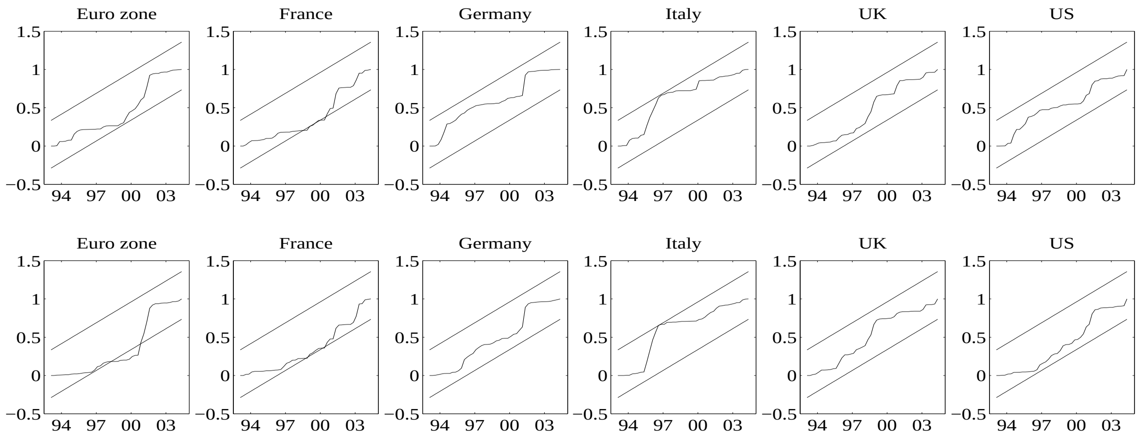


Figure 2.4: CUSUM of squares of the forward-looking Phillips curve: RULC (upper panel) and GAP (lower panel).

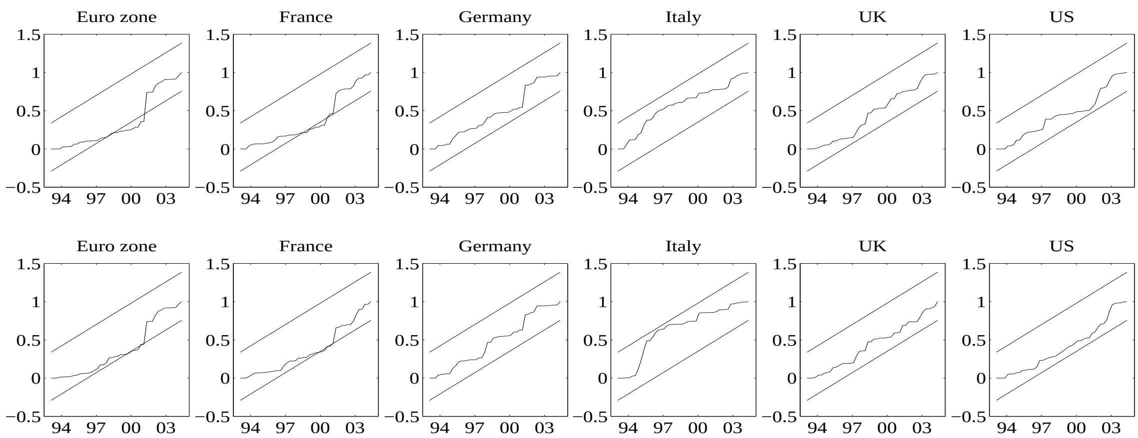


Figure 2.5: CUSUM of squares of HNKPC: RULC (upper panel) and GAP (lower panel).

2.F Discussion of the Results: Tables

	γ_f	γ_b	λ	mc	inflation expectations	sample	source
Germany	0.43	0.57	0.03	output gap	GfK Consumer Survey	1986:1-1996:4	Reckwerth (1997)
France	-	-	-	-	-	-	-
Italy	-	-	-	-	-	-	-
Euro	0.51	0.49	0.02	RULC	OECD forecasts	1977-2003	Paloviita (2006)
zone	0.46	0.54	0.08	output gap	OECD forecasts	1977-2003	Paloviita (2006)
UK	-	-	-	-	-	-	-
US	0.35	0.53	0.08	RULC	Livingston Survey	1968:4-2000:1	Adam and Padula (2003)
	0.36	0.63	0.04	output gap	Livingston Survey	1968:4-2000:1	Adam and Padula (2003)
	0.32	0.72	0.14	output gap	SPF*	1968:4-2005:1	Zhang et al. (2006)
	0.53	0.54	0.22	output gap	Greenbook	1968:3-1999:4	Zhang et al. (2006)
	0.25	0.79	0.07	output gap	Michigan Survey	1960:1-2004:4	Zhang et al. (2006)
	0.52	0.40	0.05	RULC	SPF*	1968:4-2004:2	Nunes (2005)
	0.65	0.42	0.04	output gap	SPF*	1968:4-2004:2	Nunes (2005)

* SPF=Survey of Professional Forecasters

Notes: Zhang et al. (2006) estimated the HNKPC with various data series. The estimates reported in the table use the Congressional Budget Office's estimation of potential output for the calculation of the output gap and the implicit price deflator of gross domestic product for measuring inflation.

Table 2.12: Estimation results using the survey data approach

	γ_f	γ_b	λ	mc	sample	method*	source
Euro	0.54	0.46	0.09	output gap	1998:1-2003:2	PI	Angeloni and Ehrmann (2004)
zone	0.52	0.48	0.18	output gap	1977-1997	GMM	Smets (2003)
	0.88	0.03	0.02	RULC	1970:1-1998:2	GMM1	Galí, Gertler, and López-Salido (2001)
	0.69	0.27	0.01	RULC	1970:1-1998:2	GMM2	Galí, Gertler, and López-Salido (2001)
	0.63	0.37	0.12	output gap	1970:1-1999:4	GMM	Jondeau and LeBihan (2005)
	0.50	0.50	0.05	output gap	1970:1-1999:4	ML	Jondeau and LeBihan (2005)
	0.61	0.39	0.00	RULC	1970:1-1999:4	GMM	Jondeau and LeBihan (2005)
	0.54	0.46	0.00	RULC	1970:1-1999:4	ML	Jondeau and LeBihan (2005)
	0.63	0.38	0.02	RULC	1970:1-1998:4	GMM1	Rumler (2006)
	0.61	0.39	0.11	RULC	1970:1-1998:4	GMM2	Rumler (2006)
France	0.55	0.45	-0.04	output gap	1970:1-1999:4	GMM	Jondeau and LeBihan (2005)
	0.51	0.48	0.02	output gap	1970:1-1999:4	ML	Jondeau and LeBihan (2005)
	0.60	0.40	0.01	RULC	1970:1-1999:4	GMM	Jondeau and LeBihan (2005)
	0.54	0.46	0.00	RULC	1970:1-1999:4	ML	Jondeau and LeBihan (2005)
	0.65	0.30	0.04	RULC	1970:1-1997:1	GMM	Benigno and López-Salido (2006)
	0.85	0.10	$\diamond$	RULC	1960:1-1999:4	GMM10	Leith and Malley (2007)
	0.85	0.11	$\diamond$	RULC	1960:1-1999:4	GMM40	Leith and Malley (2007)
	0.55	0.45	0.03	RULC	1978:1-2003:2	GMM1	Rumler (2006)
	0.75	0.33	0.74	RULC	1978:1-2003:2	GMM2	Rumler (2006)

Table 2.13: Estimation results using the rational expectations approach

	γ_f	γ_b	λ	mc	sample	method*	source
Ger-	0.81	0.19	0.09	output gap	1970:1-1999:4	GMM	Jondeau and LeBihan (2005)
many	0.60	0.40	0.01	output gap	1970:1-1999:4	ML	Jondeau and LeBihan (2005)
	0.69	0.31	0.00	RULC	1970:1-1999:4	GMM	Jondeau and LeBihan (2005)
	0.61	0.39	0.00	RULC	1970:1-1999:4	ML	Jondeau and LeBihan (2005)
	0.70	0.09	0.10	RULC	1970:1-1997:1	GMM	Benigno and López-Salido (2006)
	0.66	0.21	$\diamond$	RULC	1960:1-1999:4	GMM10	Leith and Malley (2007)
	0.66	0.22	$\diamond$	RULC	1960:1-1999:4	GMM40	Leith and Malley (2007)
	0.53	0.46	0.02	RULC	1970:1-2003:2	GMM1	Rumler (2006)
	0.81	0.20	0.03	RULC	1970:1-2003:2	GMM2	Rumler (2006)
Italy	0.47	0.52	-0.01	output gap	1970:1-1999:4	GMM	Jondeau and LeBihan (2005)
	0.51	0.49	-0.01	output gap	1970:1-1999:4	ML	Jondeau and LeBihan (2005)
	0.43	0.56	0.00	RULC	1970:1-1999:4	GMM	Jondeau and LeBihan (2005)
	0.53	0.47	0.00	RULC	1970:1-1999:4	ML	Jondeau and LeBihan (2005)
	0.41	0.52	0.06	RULC	1970:1-1997:1	GMM	Benigno and López-Salido (2006)
	0.47	0.45	$\diamond$	RULC	1960:1-1999:4	GMM10	Leith and Malley (2007)
	0.47	0.46	$\diamond$	RULC	1960:1-1999:4	GMM40	Leith and Malley (2007)
	0.60	0.39	0.06	RULC	1970:1-2003:2	GMM1	Rumler (2006)
	0.78	0.21	0.47	RULC	1970:1-2003:2	GMM2	Rumler (2006)

Table 2.13 (continued): Estimation results using the rational expectations approach

	γ_f	γ_b	λ	mc	sample	method*	source
UK	0.73	0.27	-0.14	output gap	1970:1-1999:4	GMM	Jondeau and LeBihan (2005)
	0.59	0.41	-0.03	output gap	1970:1-1999:4	ML	Jondeau and LeBihan (2005)
	0.82	0.18	0.03	RULC	1970:1-1999:4	GMM	Jondeau and LeBihan (2005)
	0.59	0.41	0.00	RULC	1970:1-1999:4	ML	Jondeau and LeBihan (2005)
	0.74	0.10	$\diamond$	RULC	1960:1-1999:4	GMM10	Leith and Malley (2007)
	0.74	0.11	$\diamond$	RULC	1960:1-1999:4	GMM40	Leith and Malley (2007)
US	0.77	0.23	0.03	RULC	1960:1-1997:4	GMM1	Galí and Gertler (1999)
	0.62	0.38	0.01	RULC	1960:1-1997:4	GMM2	Galí and Gertler (1999)
	0.61	0.34	0.03	RULC	1970:1-1998:2	GMM1	Galí, Gertler, and López-Salido (2001)
	0.60	0.36	0.02	RULC	1970:1-1998:2	GMM2	Galí, Gertler, and López-Salido (2001)
	0.61	0.39	-0.02	output gap	1970:1-1999:4	GMM	Jondeau and LeBihan (2005)
	0.48	0.52	0.02	output gap	1970:1-1999:4	ML	Jondeau and LeBihan (2005)
	0.68	0.32	0.01	RULC	1970:1-1999:4	GMM	Jondeau and LeBihan (2005)
	0.53	0.47	0.01	RULC	1970:1-1999:4	ML	Jondeau and LeBihan (2005)
	0.59	0.36	$\diamond$	RULC	1960:1-1999:4	GMM10	Leith and Malley (2007)
	0.59	0.36	$\diamond$	RULC	1960:1-1999:4	GMM40	Leith and Malley (2007)

* GMM=Generalized Method of Moments, ML=Maximum Likelihood Estimation, PI=Panel Instrumental Estimation

Notes: Galí and Gertler (1999) and Galí, Gertler, and López-Salido (2001) considered two alternative specifications (GMM1 and GMM2) of the orthogonality condition. Leith and Malley (2007) modeled a time-varying slope of the Phillips curve λ_t which implies that they had to calibrate the households' price elasticity of demand, and hence the mark-up of prices over nominal marginal costs. Specifically, they assumed a mark-up of 10% and 40% (GMM10 and GMM40). For more details see these papers.

Table 2.13 (continued): Estimation results using the rational expectations approach

Chapter 3

Learning Trend Inflation – Can Signal Extraction Explain Survey Forecasts?

Abstract

It can be shown that inflation expectations and associated forecast errors are characterized by a high degree of persistence. One reason may be that forecasters cannot directly observe the inflation target pursued by the central bank and, hence, face a complicated forecasting problem. In particular, they have to infer whether the observed movement of the inflation rate is due to a permanent change of policy parameters or whether it is the result of a transient shock. Consequently, it is assumed that agents behave like econometricians who filter noisy information by estimating an unobserved components model. This constitutes the trend learning algorithm employed by the forecaster. To examine whether this is a valid assumption, I fit a simple learning model to US survey expectations. The second part contains an out–of–sample forecasting experiment which shows that learning by signal extraction matches survey measures closer than other standard models. Moreover, it turns out that a weighted average of different expectation formation processes with a prominent role for signal extraction behavior is well suited to explain survey measures of inflation expectations.

3.1 Introduction

Perfect foresight and full information about constant policy parameters are common assumptions in many macroeconomic models. However, one implication of the standard New Keynesian model with rational expectations is that an unanticipated change in the inflation target would lead to a sudden jump in the level of inflation expectations. Consequently, if the model is purely forward–looking, disinflation is not accompanied by an output loss. However, both implications, the jump behavior of expectations and the absence of disinflation cost can certainly be doubted. Focussing on inflation expectations here, these can be shown to exhibit a series of features that are inconsistent with the assumption of rational expectation formation in the sense of Muth (1961). For instance, Evans and Wachtel (1992) find that U.S. inflation expectations are biased and inefficient predictors of future inflation.[1] In particular, forecast errors are found to be persistent and can be explained ex–post.

In their paper, Evans and Wachtel (1992) emphasize the empirical relevance of information constraints for the formation of inflation expectations. They state that the forecasts generated by their univariate regime–switching model exhibit some important properties of survey data on inflation expectations. A more recent series of papers also relaxes the assumption of perfect knowledge. Among others, Kozicki and Tinsley (2005), Andolfatto and Gomme (2003), Nunes (2004) and Erceg and Levin (2003) emphasize the importance of the persistence of expectations for the inflation process. One of their findings is that learning behavior is important to explain the transition dynamics of monetary policy which has implications for the design of monetary policy.[2] They also stress that this has resulted in quantitatively important welfare effects on output and interest rates during disinflation episodes. In these papers, inflation persistence stems from the fact that rational agents face a complicated forecasting problem when forming expectations. Theoretically, one reason is that they only observe noisy information, which constitutes a signal extraction problem. Like in Cukierman and Meltzer (1986), the problem arises from the fact that it cannot be distinguished between permanent target shocks and transitory shocks to the policy rate. Consequently, if the conduct of monetary policy changes over time, thereby changing the inflation target, agents face a complicated forecasting problem: the decomposition of inflation into trends and transitory components. The investigation of trend breaks in measures of inflation expectations has rarely been a subject of studies. However, the behavior of expectations following a trend shift is of importance. As noted above, sluggish expectations will constitute

[1] Among others see Roberts (1997) and Branch (2004) and the papers cited there.

[2] For instance, Andolfatto and Gomme (2003) advocate the idea that it is important for a central bank to be credible, as this significantly reduces the output inflation trade–off. For other implications of learning for an optimal monetary policy see Evans and Honkapohja (2003). Also see Cogley and Sbordone (2006) on implications for the New Keynesian Phillips curve.

the persistence of inflation rates. From the point of view of the monetary authority, it is important to know what implications a trend shift has on inflation expectations because it will determine how costly a disinflation policy will be in terms of output loss, as emphasized by Nunes (2004). Andolfatto and Gomme (2003), for example give a theoretical justification why transparency will reduce the cost of disinflation.

Another theoretical explanation for the rejection of rational expectations is that agents are boundedly rational, which means that they face resource constraints if information is costly (rational inattention) or that they lack sophistication. Within these frameworks, rational agents would, again, solve the forecasting problem by application of some (optimal) learning algorithm. These arguments have been put forward by Branch (2004) or Pajfar and Santoro (2007), among others. Moreover, Branch and Evans (2006) emphasize the importance of simple forecasting models on the part of private agents to model expectations in a setting of bounded rationality. The main finding of these papers is that aggregate expectations are the result of forecasting exercises undertaken by heterogenous agents, which are characterized by different effort and different forecasting models. In particular, these differences occur with respect to the learning rules.

Here, I basically follow the above frameworks and assume that agents have to make decisions in an environment that is characterized by noisy information. This, in turn, leads to a forecasting problem for decision makers which is more complicated than in perfect foresight models: How can the action of a central bank be interpreted in the light of new information? Does a shift of the inflation rate stem from a temporary policy action or some other temporary shock or does it reflect a permanent change in the policy parameters – i.e. the inflation target? Certainly, agents have to form expectations about these issues which will ultimately be reflected in their projections of the inflation rate. A lack of information in this context arises for different reasons. Either the monetary authority refrains from complete disclosure of the policy making process (intransparency) or the announced inflation target may not be fully credible. Moreover, to account for bounded rationality, I will assume a *simple* forecasting model on part of private agents. Following the approach of Branch and Evans (2006) and Dossche and Everaert (2005), I will restrict the analysis to univariate forecasting models. The advantage is that no assumptions about the structural relationships between variables have to be made and results will not depend on a specific theoretical model.

In the literature on learning, it has become standard to assume that agents act like econometricians who estimate the unknown parameters of the forecasting model from past data. As outlined above, here, they are also confronted with a signal extraction problem they solve by estimating permanent and transitory components of inflation in order to come up with a forecast. In particular, the forecasters will update trend perceptions each time a new observation becomes available. To do so in an efficient way, they build up an unobserved components model and make use of the Kalman filter recursions.

In chapter 3.2 I will have a closer look at the different survey measures for inflation expectations. Thereby I will briefly update and review some of the inherent characteristics of inflation expectations. Chapter 3.3 investigates whether a model of signal extraction can be fit to several survey measures for U.S. inflation expectations, thereby answering the question: Do agents update trend expectations in the light of past forecast errors? At this point, it will also be of interest to know how long it takes until agents learn about a new monetary policy regime. Taking an out–of–sample perspective in chapter 3.4, I will analyze in a forecasting experiment which forecasting model is able to approximate the respective survey closest. As suggested by the theory on heterogeneous expectations, aggregate measures of inflation expectations should be seen as a weighted average of different forecasting schemes. Hence, chapter 3.5 will be devoted to the questions which part of the survey participants learn by signal extraction and does the composition of aggregate expectations change when different periods are considered?

3.2 A First Look at Inflation Expectations

In the following section, some of the inherent features of inflation expectations and related forecast errors are explored for a selection of US surveys: the *Survey of Professional Forecasters* (SPF), the *Livingston Survey* (LIV) and the *Michigan Household Survey* (MHS). I will further investigate if the common notion of rational expectations in the way it has been introduced by Muth (1961) applies to survey measures of inflation expectations of experts and households.

3.2.1 Data Description

On the whole, I will focus on five different questions asked in the surveys mentioned above. The survey results, which will be labeled *SPF h=1* in the following contain the expected quarterly change of the GDP deflator one quarter ahead. Here, data is available from 1968 fourth quarter and ends in 2007 second quarter. *SPF h=4* gives information on the expected average change of the quarterly GDP deflator during the next four quarters. The dataset starts in 1970 second quarter and ends in 2007 second quarter. Note, that these forecasts are overlapping as the survey is conducted on a quarterly frequency. *LIV h=1* contains expectations of the annualized six month consumer price inflation six months ahead. This constitutes no overlapping forecasts as *LIV* is conducted biannually. In contrast to *LIV h=2*, which gives expectations of 12 month CPI inflation one year ahead and where the overlap is one period. I use expectations from the post war period beginning in January 1953 up to June 2007. The last survey measure of inflation expectations is given by *MHS h=12* where households are asked the following question:

> A: During the next 12 months, do you think that prices in general will go up, or go down, or stay where they are now?
> B: By about what percent do you expect prices to go (up/down) on the average, during the next 12 months?

This entails an overlap of 11 periods. *MHS* runs from January 1978 up to June 2007. As a consequence of the variety of questions under consideration, the reference series are quite different for the respective survey not only as far as the measure of price increase – and the associated variability of the series – is concerned but also with respect to the forecasting horizon.

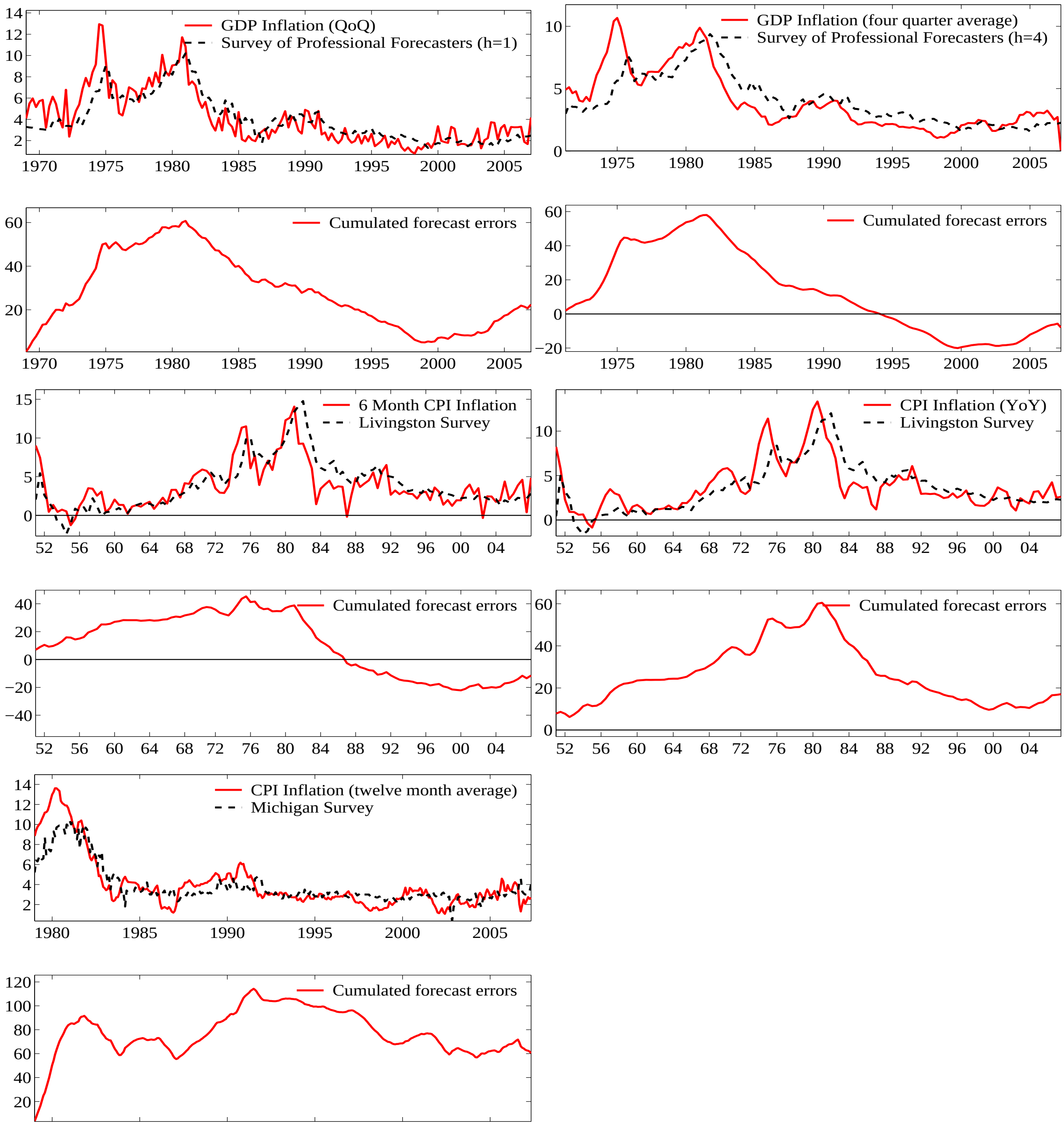
3.2.2 (Un)biasedness and (In)efficiency of Forecasts?

Define the survey expectation error as the difference between realized inflation and survey expectation with a forecast horizon of h periods. ($Error = \pi_t - \pi^e_{t|t-h}$). Thus, negative values result when the inflation rate is overestimated. Figure 3.1 visualizes the data by showing inflation expectations of *SPF*, *LIV* (beginning in 1950) and *MHS* respectively. All surveys cover the period of high inflation beginning in the seventies, reaching a peak around 1980 and falling again in the subsequent period of disinflation under the Volcker regime. They also contain the rather tranquil period of the presidency of Greenspan starting in November 1987 and the recent period under the presidency of Bernanke since 2006. It becomes clear that cumulated forecast errors tend to follow the pattern of the inflation rate itself. This means that during phases of rising inflation like in the 1970s a repeated underprediction of inflation rates can be observed. As inflation comes down to moderate levels in due course, the cumulated forecast error decreases again most notably for *SPF* and *LIV*, which means that inflation is overpredicted during that period. Also note that in almost all cases considered here, the cumulated forecast error displays strong persistence. This means that an error in one period is not completely offset in the subsequent period but agents are sluggish when changing expectations. Thus, there seems to be a case for bounded rationality.

The recent findings can also be investigated more formally. In the following, I basically update some of the results on survey expectations found in Evans and Wachtel (1992) whose sample ends in 1991. Following the rational expectations hypothesis of Muth (1961), forecast errors as defined above should follow a zero mean white noise process if survey participants form rational expectations. This requires expectations to be unbiased and efficient in the sense that no information is omitted when forming expectations. To check if unbiasedness is a valid assumption, I run the regression described in equation (3.1) and test if $a = 0$ and $b = 1$ by means of a Wald test.

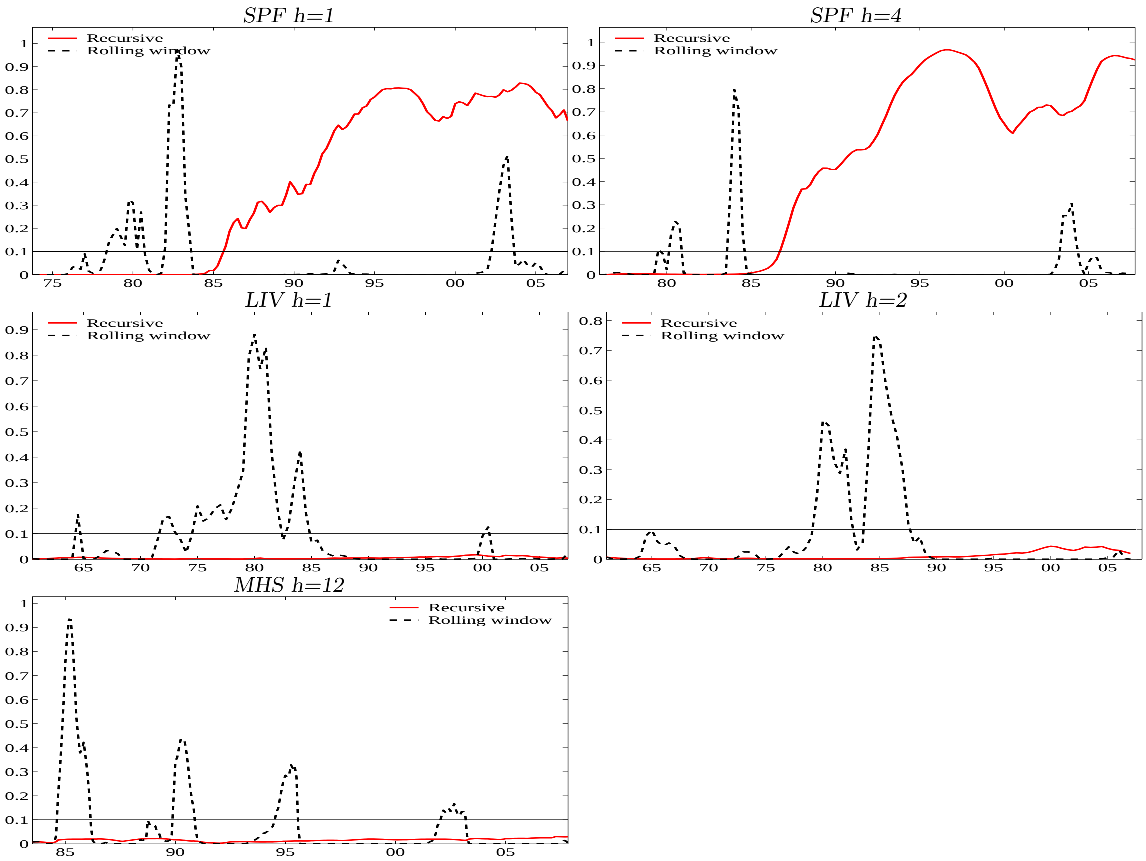
$$\pi_t = a + b\pi^e_{t|t-h} + \epsilon_t \tag{3.1}$$

Here, $\pi^e_{t|t-h}$ is the expected inflation rate conditional on the information set at time $t - h$. Figure 3.2 plots recursive Wald tests, as well as a test based on a rolling window of five years for *SPF* and *MHS* and ten years for *LIV*. Considering the whole sample, unbiasedness is not rejected for both questions asked in SPF. If sub–samples are considered by recursive estimation, the SPF provides biased expectations up to the mid–eighties when there has been the large swing in the inflation rate. Afterwards – with the period of disinflation having passed – the test indicates that expectations are unbiased. Rolling window estimates point into the same direction. The finding can also be confirmed by looking at the cumulated forecast error which returns to zero in the mid–nineties, thereby indicating that – on average – expectations have been unbiased.



Note: The first panel shows the expected annualized quarterly GDP inflation one quarter ahead along with the realizations. The sample begins in 1968 Q4 and ends in 2007 Q2. The second panel depicts the four quarter moving average of GDP inflation along with expected average annualized inflation from *SPF* during the next $h = 4$ quarters where the sample runs from 1970 Q2 to 2007 Q2. The third panel depicts the annualized six months growth rate of CPI along with expected inflation with a forecasting horizon of $h = 1$ half years. The fourth panel contains the one year growth rate of prices along with expected inflation with a forecasting horizon of $h = 2$. Both measures are taken from *LIV* where the sample runs from 1950 I to 2007 I. The last panel shows CPI Inflation as the twelve months moving average growth rate of prices along with expected CPI inflation with a forecasting horizon of $h = 12$ months from *MHS*. The sample runs from 1978 M1 to 2007 M6. The lower part of each panel shows a plot of the cumulated forecast errors $\sum_{\tau=0}^{t}(\pi_\tau - \pi^e_{\tau|\tau-h})$ up to time t.

Figure 3.1: Inflation expectations from MHS, SPF, LIV

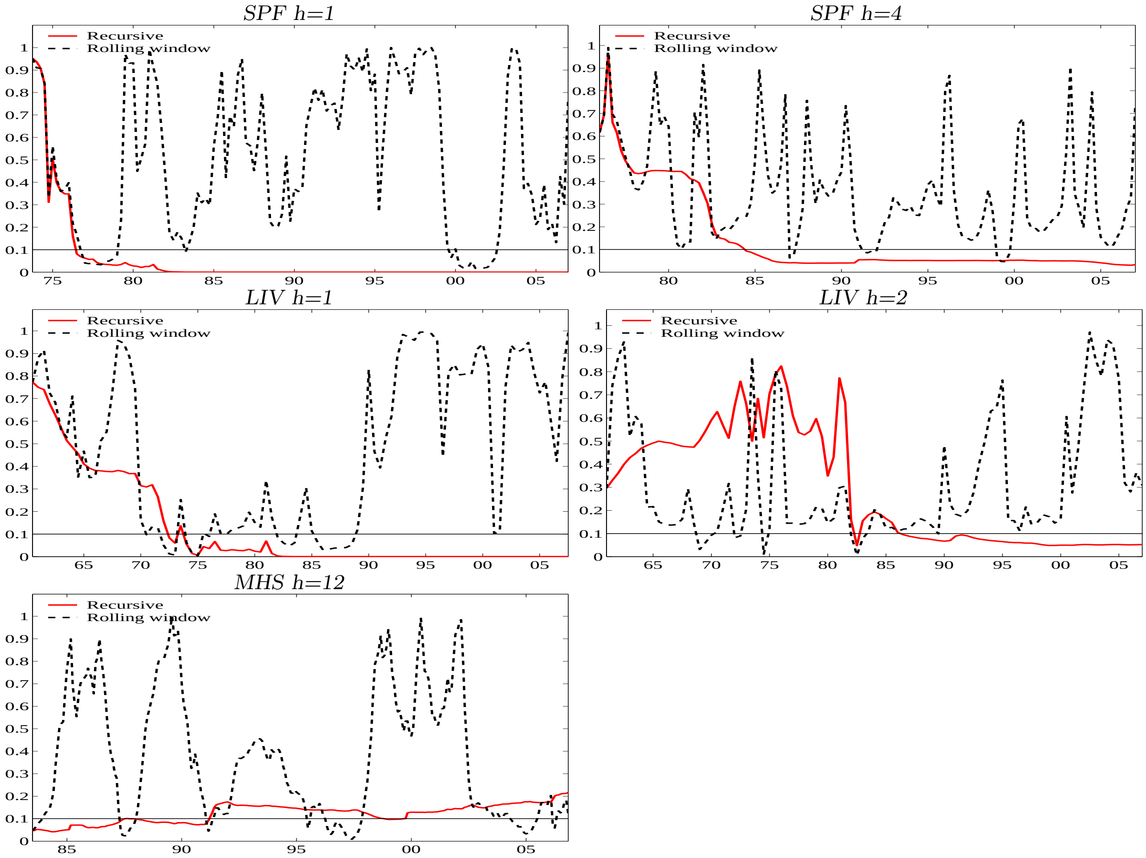


Note: The solid red line shows p–values for a recursive Wald–test of $H_0 : a = 0, b = 1$. The dashed line represents p–values based on a rolling window. The initial estimation period and the rolling window cover 5 years for *SPF*, 10 years for *LIV* and 5 years for *MHS*. The sample runs from 1968 Q4 - 2007 Q2 (*SPF h=1*), 1970 Q2 - 2007 Q2 (*SPF h=4*), 1950 I - 2007 I (*LIV h=1* and *LIV h=2*) and 1978 M1 - 2007 M6 (*MHS*).

Figure 3.2: Recursive Wald–test *SPF, LIV, MHS*

On the other hand, LIV, which is questioned on a semiannual frequency, is clearly biased. But when estimated on a rolling window beginning in the late seventies, which does not cover much of the period of high inflation, it turns out to be unbiased. The MHS is biased throughout the whole sample, whereas the rolling window tests indicate unbiasedness from time to time – especially during the mid–eighties again. Keeping in mind that it is an household survey and that the sample does not cover all of the high inflation period either, this does not come a surprise. When compared to the representation of the cumulated forecast error, biasedness is confirmed by the fact that the zero line is not crossed although the cumulated errors clearly stabilize in the second half of the sample. Thus, whether an expectations series is biased crucially hinges on the time period considered. One conclusion which can be drawn here is that biasedness of expectations seems to be a small–sample problem in the sense that samples are finite.

Another way to test if survey participants tend to adjust their beliefs in the light of new information in a sluggish manner, is to directly look at autocorrelation. If forecast errors are highly persistent, then – after the concept of rational expectations of Muth (1961) – one concludes that forecasts are formed inefficiently. Note, however, that for overlapping forecasts, a shock that occurs within the forecasting period cannot be taken into account by the forecaster and the same mistake is necessarily repeated. Therefore, I present a test proposed by Cumby and Huizinga (1992) which allows to deal with the fact that forecasts are $h-1$ dependent. Figure 3.3 shows recursive p–values for a test of first–order autocorrelation of forecast errors based on the ℓ–statistic of Cumby and Huizinga (1992), as well as p–values based on a rolling window of five years for *SPF* and *MHS* and ten years for *LIV*.



Note: The solid lines represent recursive p–values for a test of autocorrelation of forecast errors based on the ℓ–statistic proposed by Cumby and Huizinga (1992). The dashed lines give p–values for the ℓ–statistic for a rolling window. The initial estimation period and the rolling window covers 5 years for *MHS* and *SPF* and 10 years for *LIV*. The sample runs from 1968 Q4 - 2007 Q2 (*SPF h=1*), 1970 Q2 - 2007 Q2 (*SPF h=4*), 1950 I - 2007 I (*LIV h=1 and h=2*) and 1978 M1 - 2007 M6 (*MHS*).

Figure 3.3: Persistence of forecast errors of *SPF*, *LIV*, *MHS*

For the whole sample, indeed, the ℓ–statistic is significant for *SPF* and *LIV*. However, there seems to be a period in the beginning of each sample where no persistence of forecast errors can be found. Looking at *SPF* and *LIV* it becomes clear that autocorrelation is found from the late seventies or the early eighties on. Furthermore, considering the rolling window estimates for *LIV* in the eighties – a period of large swings of inflation – the test indicates that forecast errors are persistent during that period of time. As far as *MHS* is considered, errors exhibit significant persistence during the eighties. Based on a recursive scheme, this is not confirmed for the whole sample. But keep in mind that the sample starts only in 1978 and does not contain the whole period of high inflation. If persistence of errors is a major problem during periods of large swings, then – in terms of sample size – estimates based on the whole sample are probably dominated by the longer period of only moderate movements in case of *MHS*. In order to investigate this in more depth, I also analyze if forecast errors are larger and tend to exhibit more persistence when the underlying variable experiences large changes. The correlation between forecast errors and inflation is presented in the left part of table 3.1, whereas the right part gives the correlation with forecast changes.

	Cross correlation of forecast error							
	$\Delta\pi_t$				$\Delta\pi^e_{t\|t-h}$			
$\pi_t - \pi^e_{t\|t-h}$	lag_0	lag_1	lag_2	lag_4	lag_0	lag_1	lag_2	lag_4
SPF h=1	0.54	0.27	0.03	0.15	0.43	0.17	0.10	0.20
SPF h=4	0.73	0.72	0.66	0.44	0.76	0.78	0.74	0.51
LIV h=1	0.64	0.27	0.19	−0.04	0.55	0.20	0.14	−0.05
LIV h=2	0.77	0.59	0.33	−0.03	0.70	0.60	0.33	0.03
MHS h=12	0.76	0.73	0.66	0.56	0.61	0.55	0.53	0.51

Note: The sample of survey forecast errors runs from $1969Q1 - 2007Q2$ (*SPF h=1*), $1971Q2 - 2007Q2$ (*SPF h=4*) and $1950\,II - 2007\,I$ (*LIV h=1*) and $1951\,I - 2007\,I$ (*LIV h=2*) and $1979M1 - 2007M6$ (*MHS*). The displayed lag lengths coincide with very different time intervals. Due to the different frequencies of the surveys, four lags imply for the *MHS* 4 months, for the *SPF* one year and for the *LIV* two years.

Table 3.1: Cross correlation of forecast errors

Errors are apparently positively correlated with the change of the inflation rate. This is compatible with the view that an overestimation of inflation comes along whenever the inflation rate declines or has declined the period before. In other words, the higher the decline, the larger is the associated overestimation which implies that forecasters do not respond very rapidly to shocks in the inflation rate. This is in line with the findings from figure 3.3. Interestingly, the same result can also be found for the correlation of forecast errors with forecast changes. Whenever an underprediction occurs, there is a tendency to raise forecasts in subsequent periods. In general, this shows that forecasts do not respond very quickly to past errors.

To conclude, expectations are formed in a way inconsistent with the common concept of rationality which relies on full information and perfect foresight. Consequently, a number of studies have also come to the conclusion that rational expectations do not provide a good description of expectation formation processes[3]. It is important to note that expectational errors are found to be persistent especially in periods of large inflation movements. Moreover, for *SPF*, a bias is only found for such a period but not for the whole sample, whereas the other surveys are unbiased in a sub–sample around the mid–eighties. From table 3.1 it can also be inferred that forecast errors are larger during periods which are characterized by large swings of inflation and where forecasting is more complicated. Note, that all surveys seem to behave very similar with respect to bias and persistence of expectational errors despite the fact that respondents and reference variables as well as forecasting horizons differ considerably. In the following, I will investigate whether these characteristics are related to signal–extraction in a noisy environment.

3.3 Learning With a Simple Forecasting Model

3.3.1 Motivation and General Framework

Standard New Keynesian models, which assume rational expectations, will predict that unanticipated changes in the inflation target of a central bank will lead to an immediate jump of expectations and the level of inflation. As shown in section 3.2.2 this is not a realistic assumption because expectations are characterized by significant inherent persistence. Moreover, section 3.2.2 suggests that, although inflation may move rather quickly in disinflation episodes, expectations adjust only sluggishly. In Andolfatto and Gomme (2003), for instance, the authors argue that these features are observed not because agents are ignorant. They simply face a complicated forecasting problem which introduces sluggish adjustment to a new target inflation. Thus, disinflation comes along with significant output loss. The key assumption of the analysis in the present paper is that private decision makers cannot directly observe the inflation target pursued by the central bank. This may be due to the fact that the central bank is not transparent or that it has low credibility. Moreover, there may be information problems as far as the timing of the change is concerned. Note, that this situation is well applicable for the case of U.S. inflation during the eighties. Moreover, it may be valid even today as the FED does not announce an explicit target rate.

Following Andolfatto, Hendry, and Moran (2002), Dossche and Everaert (2005), Kozicki and Tinsley (2005) and Erceg and Levin (2003), I assume that agents solve a signal extraction problem in order to infer whether the observed movement of infla-

[3] See for example Roberts (1997) and the papers cited there.

tion is due to a transient monetary policy shock or whether the monetary authority has changed its inflation target, which will result in a permanent change of the level of inflation. In particular, they have to form expectations about trend inflation, which, in these models, coincides with the central bank's inflation target.

More intuitively, the signal extraction problem can be described as follows. In the first place, private decision makers observe an interest rate that is higher than would be the outcome of a strict application of the interest rule. They perceive the action of the central bank as if there had been a discretionary transitory shock to the system. However, it will be offset by application of the monetary policy rule in due course when the monetary authority wants to achieve its target again. Consequently, after having observed a monetary tightening, agents expect that, in the next quarter, inflation will rise to target levels again. If, however, inflation remains at lower than target levels for more a longer period of time, then decision makers will conclude that the target has changed. But they are not informed about the magnitude of the change. This constitutes the signal extraction problem as they cannot observe to which extent the movement of the inflation rate is due to a temporary monetary action and which part of the change in inflation rates is the result of a permanent shift. Thus, they have to form expectations about the target that is pursued by the central bank. Even if they do so in an optimal way, it will take some time until they can infer the correct new target from the observed action. Consequently, expectational errors will occur although decision makers are not ignorant. A fact that has been emphasized by Andolfatto, Hendry, and Moran (2006) and Evans and Wachtel (1992).

One obvious solution to the signal extraction problem is given by application of the Kalman filter. Due to its recursive formulation, it provides the theoretical concept describing how private agents learn about trend inflation in this study. Moreover, this learning rule is optimal, provided that agents know the true model of the economy. However, as for instance argued in Branch and Evans (2006), the concept of bounded rationality is more appropriate here. According to the concept, agents do not know the true model of the economy in reality but they will rather apply a simple model that is easily applicable and serves the respective purpose. For instance, resource constraints or limited tools for information processing will advocate the use of a simple model on behalf of private agents.

Moreover, agents may be quite heterogeneous with respect to resource constraints or as far as the evaluation of a bad forecast is concerned. Therefore, the theory of heterogeneous expectations outlined in – among others – Branch (2004) suggests that each agent undertakes a different effort to find a suitable forecasting model and to get an estimate of the model's parameters. That is one reason why one part of the agents may, for instance, be rational or trend learner, another part will simply be backward–looking or completely ignorant.

This is the framework for the present analysis, where, on the one hand, survey participants are characterized by a lack of information with respect to target inflation. On the other hand, they will rely on a simple forecasting model. In particular, they do not make the effort to find the true model of the economy. Moreover, it may be the case that more than one "type" of forecaster contributes to the survey result.

3.3.2 The Forecasting Model

In the following, I will formalize expectation formation of private agents. In the spirit of bounded rationality, I assume that agents form expectations according to a simple forecasting model in order to deal with limited information processing capacities. The basic model consists of three (unobserved) components. It comprises a time–varying trend $\overline{\pi}_t$ that captures the permanent component of inflation and, in addition, inflation inherits cyclical movements $\widehat{\pi}_t$ and unsystematic shocks ϵ_t. The model is given by equations (3.2) to (3.4) which constitute the data generating process for inflation expectations.

$$\pi_t = \overline{\pi}_t + \widehat{\pi}_t + d_t + x_t + \epsilon_t \qquad \epsilon_t \sim N(0, \sigma_\epsilon^2) \tag{3.2}$$

$$\overline{\pi}_{t+1} = \overline{\pi}_t + \eta_t \qquad \eta_t \sim N(0, \sigma_\eta^2) \tag{3.3}$$

$$\begin{pmatrix} \widehat{\pi}_{t+1} \\ \widehat{\pi}^*_{t+1} \end{pmatrix} = \rho \begin{pmatrix} \cos\lambda & \sin\lambda \\ -\sin\lambda & \cos\lambda \end{pmatrix} \begin{pmatrix} \widehat{\pi}_t \\ \widehat{\pi}^*_t \end{pmatrix} + \begin{pmatrix} \kappa_{t+1} \\ \kappa^*_{t+1} \end{pmatrix} \tag{3.4}$$

$$\begin{pmatrix} \kappa_{t+1} \\ \kappa^*_{t+1} \end{pmatrix} \sim N(0, \begin{pmatrix} \sigma_\kappa^2 & 0 \\ 0 & \sigma_\kappa^{2*} \end{pmatrix})$$

This is basically the widely used *trend plus cycle model* introduced by Harvey (1989) augmented by a vector of dummy variables d_t and possibly some exogenous variables x_t, that may be useful when forecasting inflation. Above all, the model accounts for the fact that there is a signal extraction problem as the distinct components are not observable but have to be estimated. When learning about the unobserved target, forecasters are assumed to update their perception each time a new observation becomes available. To be more precise, they learn from their forecast errors of the past. Like in Erceg and Levin (2003) agents make use of a so–called constant gain learning rule. This means that, after having observed a forecast error, each component is updated by a constant part of the error. Technically speaking, when estimating the unobserved components, agents learn from noisy information contained in the forecast errors $\nu_t = \pi_t - E_{t-h}\pi_t$. In particular, there is learning of both, trends as well as cycles at the same time. This is in contrast to the standard adaptive learning scheme like in Pajfar and Santoro (2007), for example.

Consequently, forecasters will behave very much like econometricians who estimate an unobserved components model. In order to solve the signal extraction problem optimally, they can make use of the Kalman filter recursions to obtain an estimate of the unobserved components. The optimal gain is then given by the so–called Kalman gain[4].

3.3.3 Fitting Survey Expectations

The Framework

The present section tests whether expectation formation implied by equations (3.2) to (3.4) can explain survey measures of expectation. In the following, an in–sample perspective is taken. However, if the data generating process can be described by the above equations, the process for inflation expectations can be written down in the form of a state–space model which consists of an observation equation (3.5) and state equations (3.6) to (3.8), which describe how unobserved components are estimated.

$$\pi_{t+1|t} = \overline{\pi}_{t+1|t} + \widehat{\pi}_{t+1|t} + \beta_{t+1} d_{t+1} \tag{3.5}$$

The subscript $t|t-1$ denotes the mean of the distribution at t predicted from information up to time $t-1$. The Kalman filter recursions, which are employed to estimate the unobserved components are reformulated such that they only contain predicted state variables. Expectations of trend and cycle in the next period (based on the last prediction) are given by:

$$\begin{aligned}
\overline{\pi}_{t+1|t} &= \overline{\pi}_{t|t-1} + \mathcal{K}_{1,t}\nu_t && (3.6)\\
\widehat{\pi}_{t+1|t} &= \rho\cos\lambda\widehat{\pi}_{t|t-1} + \rho\sin\lambda\widehat{\pi}^*_{t|t-1} + \widetilde{\mathcal{K}}_{2,t}\nu_t && (3.7)\\
\widehat{\pi}^*_{t+1|t} &= -\rho\sin\lambda\widehat{\pi}_{t|t-1} + \rho\cos\lambda\widehat{\pi}^*_{t|t-1} + \widetilde{\mathcal{K}}_{3,t}\nu_t && (3.8)
\end{aligned}$$

Here, $\nu_t = \pi_t - \pi_{t|t-1}$ denotes the expectation error of the last period. It can be shown that $\widetilde{\mathcal{K}}_{2,t} = \rho\cos\lambda\mathcal{K}_{2,t} + \rho\sin\lambda\mathcal{K}_{3,t}$ and $\widetilde{\mathcal{K}}_{3,t} = -\rho\sin\lambda\mathcal{K}_{2,t} + \rho\cos\lambda\mathcal{K}_{3,t}$ where $\mathcal{K}_{i,t}$ represents the gain parameter according to which unobserved components are

[4]See Harvey (1989), chapter 3.2 or Hamilton (1994), chapter 13.2 for a derivation of the Kalman updating algorithm which yields the conditional mean of the distribution of unobserved components. The procedure minimizes the squared forecast errors provided the system is linear and disturbances are Gaussian white noise processes.

updated when a misperception of inflation occurs[5]. In particular, $\mathcal{K}_{1,t}$ determines the update of the estimated trend and $\mathcal{K}_{2,t}$ determines the updating scheme with respect to the transitory part. $\mathcal{K}_{3,t}$ captures an indirect effect of misperceptions on the update of the transitory component and is given for completeness. The optimal forecasting scheme with respect to the data generating process given in section 3.3.2 is given by the Kalman filtering rule. Thus, trend expectations should be updated by an amount equal to the implied Kalman gain.

In order to investigate the properties of the survey, the conditional mean in the future $\pi_{t+h|t}$ is replaced by survey expectations $\pi^e_{t+h|t}$. The forecast error ν_t is exchanged with its observed counterpart, associated with the respective survey with forecast horizon h, i.e. $\nu^e_t = \pi_t - \pi^e_{t|t-h}$[6]. For estimation purpose, ε_t reflects the part of survey expectations which is not explained by the model.

In this simple univariate setting, equations (3.7) and (3.8) capture the persistence of the transitory part of expectations. In addition, the model allows for signal extraction – i.e. learning from repeated forecast errors, because trend learning is necessarily associated to the estimation of the cyclical component. Hence, a part of the forecast error is related to misperceptions of the cyclical part. Thus, we obtain in–sample estimates for the gain parameters $\mathcal{K}_{1,t}$ to $\mathcal{K}_{3,t}$ for each survey. These determine the speed of learning of survey participants when a change of unobserved trend inflation occurs. Furthermore, we can test if extracted trend expectations are characterized by persistence and it is possible to infer the speed of trend learning.

Estimation Results

The system which consists of equations (3.5) to (3.8) is estimated by maximum likelihood. The diffuse likelihood is computed by the Kalman filter with diffuse prior density of the initial state vector. The parameter vector $\psi = \begin{pmatrix}\sigma^2_\varepsilon & \lambda & \rho\end{pmatrix}$ consists of the variance of the irregular component ε_t, the cycle length λ and the so–called dampening factor ρ. It is reparameterized such that the theoretical restrictions are fulfilled (see appendix 3.A for details). Dummy variables are set when the outlier test proposed by Harvey and Koopman (1992) indicated an outlier. Common regression diagnostics and a histogram of past forecast errors ν_t are given in appendix 3.C. The gain parameters $\mathcal{K}_{i,T}$ can be extracted from the smoothed state vector and are not restricted during estimation, whereas the smoothing recursions also yield estimated standard errors. The estimated parameters are summarized in table 3.2

[5] Note that gain parameters $\mathcal{K}_{i,t}$ relate to reduced form parameters $\widetilde{\mathcal{K}}_{i,t}$ in a linear way.

[6] In the cases where $h > 1$, this implies that forecasters apply some kind of *direct* multi–step forecasting. Hence, the gain parameters cannot be interpreted as the usual Kalman gains any more. In other words, signal extraction with the Kalman filter would only yield a minimum forecast error, if it relies on the one–step–ahead forecasting error from *SPF h=1* and *LIV h=1*.

for each survey. The in–sample observation period generally runs from 1972 to 2007 for *SPF* and *LIV*. For *MHS* the sample only begins in 1979, as the survey has not been published before. Hence, there is at least one possible structural break agents may have learned which is commonly associated with the beginning of the Volcker era. Taking a look at equations (3.5) to (3.8), it becomes clear that there is only one error term in the system (ε_t) which captures irregularities. Hence, in a technical sense, the estimated unobserved components are non–stochastic as far as the Kalman recursions are concerned – i.e. the dynamics of all dependent variables is solely explained by past forecast errors and autoregressive elements.

		$\widehat{\sigma}^2_\varepsilon$	$\widehat{\lambda}$	$\widehat{\rho}$	$\mathcal{K}_1$	$\mathcal{K}_2$	$\mathcal{K}_3$
SPF h=1	72–07	0.469 (0.690)	0.261 (0.244)	0.953 (0.139)	0.119 (0.004)	0.010 (0.018)	0.066 (0.013)
SPF h=4	72–07	0.218 (0.318)	0.006 (0.611)	0.002 (0.014)	0.077 (0.002)	317.5 (–)	1.041 (–)
LIV h=1	72–07	2.020 (3.253)	2.307 (–)	0.000 (0.000)	0.136 (0.009)	0.016 (–)	0.063 (–)
LIV h=2	72–07	0.660 (0.968)	0.000 (0.622)	0.001 (0.010)	0.160 (0.007)	315.0 (–)	0.008 (–)
MHS h=12	79–07	0.171 (0.246)	0.001 (0.013)	0.925 (0.080)	−0.004 (0.002)	0.057 (0.004)	−0.602 (0.234)

Note: The table shows maximum likelihood estimates of the parameters and estimated constant gain coefficients of the surveys. Numbers in parenthesis are standard errors. A (–) indicates that numerical estimates are not available, which is the case if ρ is close to the lower bound of zero. Estimated parameters are presented for completeness.

Table 3.2: Estimated parameters

Turning to table 3.2, one observes that estimated variances σ^2_ε, which are presented in the first column, are in a plausible range. However, estimates are not very precise. The estimated cycle parameters, which are given in column two and three, show a more heterogeneous picture. In the case of *SPF h=1* the estimate is 0.26 which implies a cycle length of approximately 24 quarters. $\widehat{\rho}$ which determines the "sluggishness" of the cycle is in a plausible range being compatible with the concept of an autoregressive transitory component. Turning to the longer forecasting horizon (*SPF h=4*), cycle length is considerably longer. Interestingly, $\widehat{\rho}$ is zero, which implies that there is no autoregressive part in the system and the cyclical dynamics can solely be explained by past forecast errors made four periods ago. As far as *LIV* is concerned, the survey which has a shorter horizon (*h=1*), yields a dampening factor of zero. Moreover, much of the variation of expectations in *LIV h=1* is captured by the error term, which has a rather high variance. Cycle parameters for *LIV h=2* also indicate that autoregressive elements cannot explain the cyclical behavior of expectations. In other words, past forecast errors seem to be responsible for most of the persistence in the cyclical component. Most notably, the model seems to be supportive for the trend learning hypothesis. The gain parameters of *SPF* and *LIV*, although not restricted during estimation, lie between 0.08 and 0.16, depending on the survey. This means that in the case of *SPF h=1*, for instance, about 12 percent of the last error is attributable to trend mis–perceptions. Moreover, in both surveys, $\mathcal{K}_1$ is highly significant which leads to the conclusion that trend updating is an important characteristic of inflation expectations.

Interestingly, estimation results differ quite a bit for MHS. Here, we observe a λ which is practically zero, meaning there is no cyclicality. However, the dampening factor is close to one and, hence, persistence of the cyclical component is captured to a large extent by autocorrelation; but also learning gains seem to be important. However, estimates of $\mathcal{K}_1$ are significantly negative, which would imply that the adjustment is made into the wrong direction. But also note, that trend adjustment is economically unimportant when compared to the adjustment of "cycle" expectations.

Trend and Cyclical Components of Expectations

Figures 3.4 to 3.8 depict the unobserved components which are extracted by the Kalman smoothing recursions beginning in 1972 and 1979, respectively.

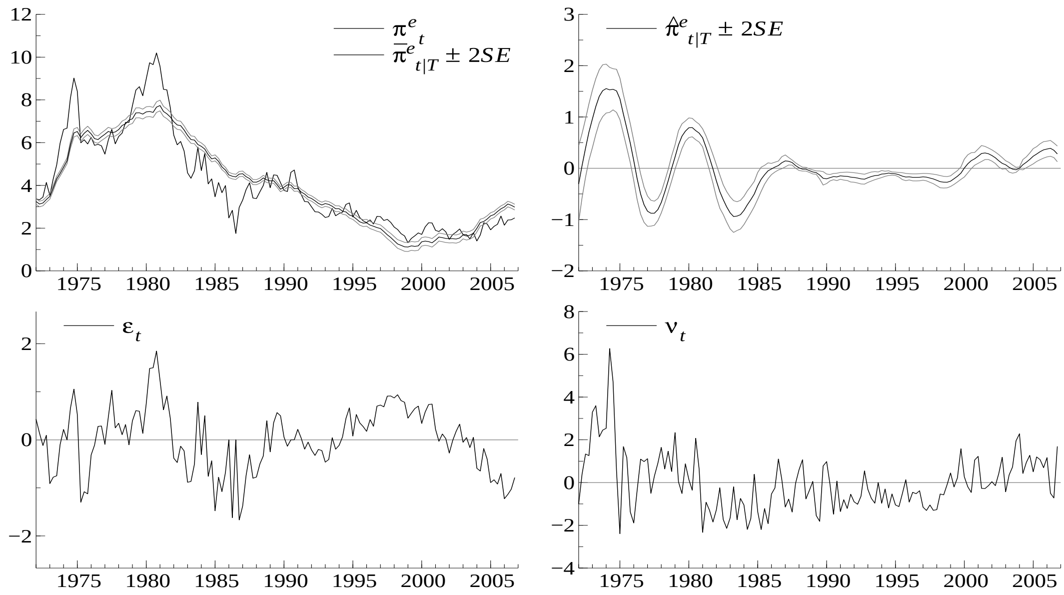
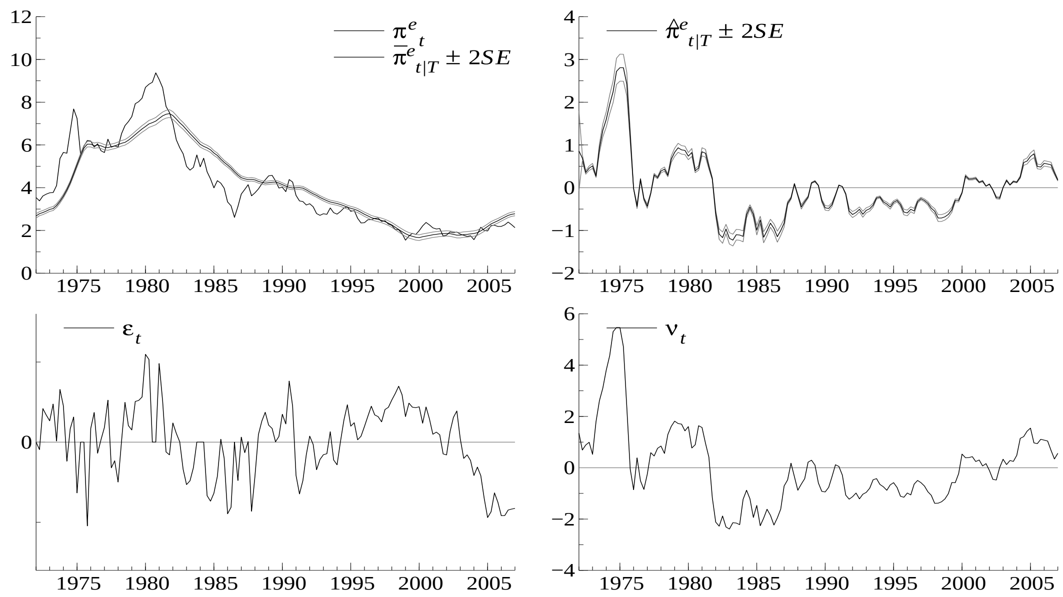


Figure 3.4: Learning model *SPF h=1*



Note: The upper left panel shows inflation expectations together with the smoothed expected trend $\pi^e_{t|T}$ where T is the last available observation. The second panel depicts the smoothed cyclical component $\widehat{\pi}^e_{t|T}$. The irregular component ε_t and the forecast error ν_t are given in the lower part.

Figure 3.5: Learning model *SPF h=4*

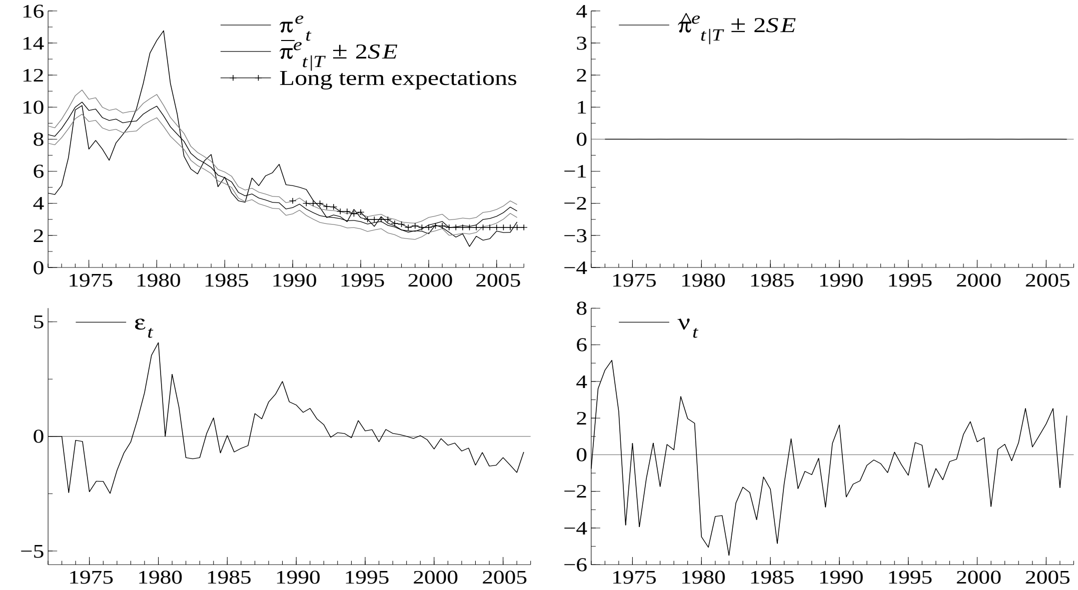
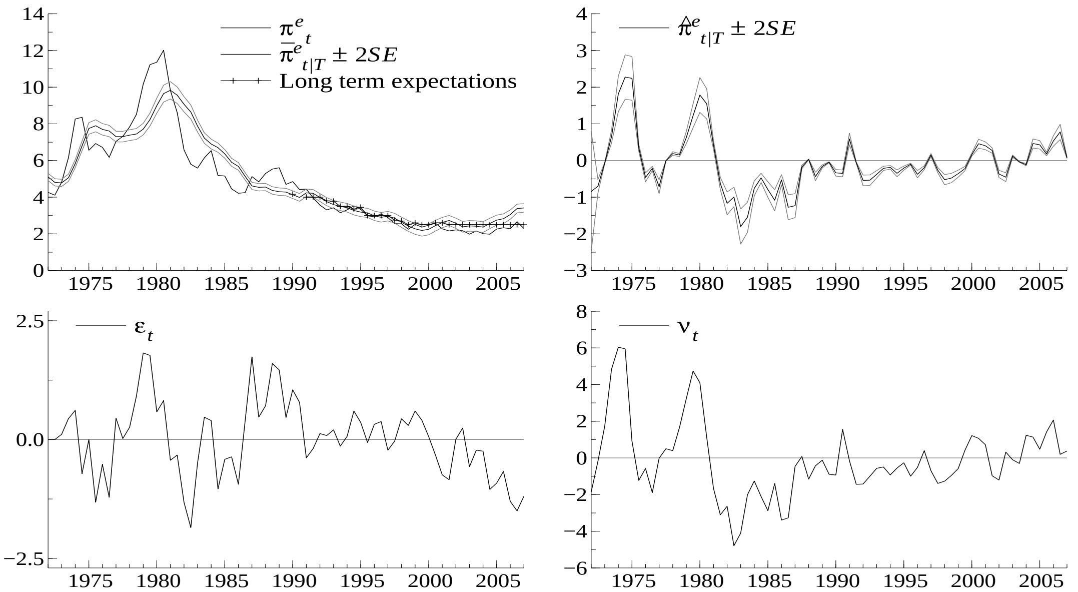
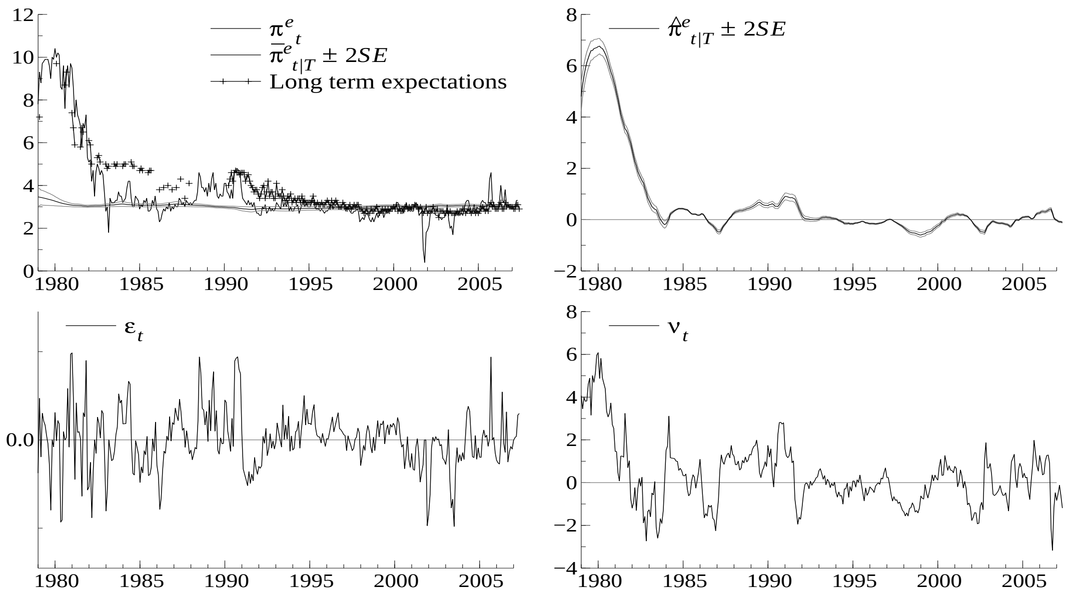


Figure 3.6: Learning model *LIV h=1*



Note: The upper left panel shows inflation expectations together with the smoothed expected trend $\pi^e_{t|T}$ where T is the last available observation. It also contains long–term (10 years) inflation expectations from LIV. The second panel depicts the smoothed cyclical component $\widehat{\pi}^e_{t|T}$. The irregular component ε_t and the forecast error ν_t are given in the lower part.

Figure 3.7: Learning model *LIV h=2*



Note: The upper left panel shows inflation expectations together with the smoothed expected trend $\pi^e_{t|T}$ where T is the last available observation. It also contains long–term (5 years) inflation expectations from MHS. The second panel depicts the smoothed cyclical component $\widehat{\pi}^e_{t|T}$. The irregular component ε_t and the forecast error ν_t are given in the lower part.

Figure 3.8: Learning model *MHS h=12*

The upper left hand graph depicts the original expectation series π^e_t together with the estimated trend component and an error band of two standard deviations. In general, it becomes apparent that trend expectations do not jump. By contrast, they are very sluggish and move rather slow. Note that, although stemming from different surveys and comprising quite different target variables, the general picture presented in the figures is quite similar. As a consequence, it takes until the mid–nineties to obtain trend expectations that are around 3%. Turning to *SPF*, this is reflected by the fact that estimated gain parameters $\mathcal{K}_1$ in table 3.2 are rather low. This implies that each quarter – by and large – only ten percent of the forecast error is attributable to trend misperceptions. This would be a possible explanation why, in the event of a changing target inflation, expectations show persistence.

Inspection of the transient component reveals that cyclical movements are much more pronounced until the mid–eighties. The same seems to be true for forecast errors. As outlined before, cyclical movements of *SPF h=4* are related to past forecast errors only. As the model is written, the irregular component of the signal equation captures all of the unexplained part of the dynamics. Admittedly, here it still displays some systematic movements. In the first place, this is a hint that assuming learning behavior is not sufficient because it cannot explain all of the expectation formation process perfectly. An observation strongly emphasized by the literature on heterogeneous expectations[7]. Secondly, however, there may also be a time–varying nature to the expectations formation process. Just imagine that during tranquil periods, like after 1987, it may be worthwhile for the agents to adapt a simple backward–looking forecasting scheme. This could, in principal be tested by splitting the sample. However, for an estimation of the structural time–series model the period is rather short. Note, however, that there seems to be a change in the behavior of ε_t, which seems to display less systematical movements during the Volcker period. That is one reason why, for the later analysis, the sample will be split in 1987. Another observation can be made when comparing trend estimates for *SPF h=1* and *SPF h=4*. The peak of trend expectations is earlier for the survey with the shorter forecasting horizon. This is explained by the fact that information on the forecast error is available earlier, which would also explain some of the differences of the cyclical component of both series.

Now turning to *LIV h=1*, we observe a flat cyclical pattern, whereas the residual component captures most of the dynamics. Hence, past forecast errors seem to explain trend dynamics but do not account for the transitory movements. However, the model seems not to be completely at odds, as it is capable to reproduce a trend that is consistent with long–term inflation expectations contained in the Livingston Survey. Moreover, the trend is learned quite slowly as estimated gain parameters imply that every half year by and large 15% of the error is used to adjust trend expectations. Compared to figure 3.6, a slightly different picture emerges for the longer forecast horizon of one year. Here, as with *SPF h=4*, we observe cyclical movements which are explained by past forecast errors. The peak is again later for a one–year forecasting horizon than for six–month expectations, which is one reason for the differences of the transient components of *LIV h=1* and *LIV h=2*. However, the trend is also consistent with the long–term expectations.

Considering *MHS h=12*, where the sample starts in 1979, we observe a different picture than before. Trend expectations estimated with the proposed learning model are flat, which indicates that there is no update of the trend due to past forecast errors. A finding that comes along with estimates of $\mathcal{K}_1$ being virtually zero. Moreover, trend expectations are inconsistent with long–term inflation expectations from *MHS*. Interestingly, the observed long–term expectations seem to lie above *MHS h=12* for most of the time. This is in particular true for the disinflation episode. Also note that the series features virtually no cyclicality after 1987. Additionally, observed forecast errors for that period of time are less systematic than in the other surveys[8]. Hence, learning may be an explanation for the dynamics of the "cyclical" component, but not for the trend. This may have different reasons.

[7] See, for instance, Branch (2004) and the papers cited there.

[8] Also compare the cumulated forecast errors presented in figure 3.1.

First, the estimations are based on a shorter sample which begins in 1979. Second, *MHS* polls households which are probably faced with different incentives and restrictions than professionals when making a forecast. Moreover, in contrast to the other surveys, the model may not give a suitable description because an update of expectations is possible more quickly than assumed by the model. As *MHS* is conducted on a monthly frequency, survey participants probably will not wait twelve periods to update their information set, which would be the case here.

Generally speaking, it is possible to fit survey data on inflation expectations to the simple learning model presented here. It produces the sluggishness of expectations in the event of shifts in target inflation. The reason is that agents are learning from a noisy signal which – in the univariate setup here – is the past forecast error. Note, that $\mathcal{K}_i$ has not been restricted during estimation. Nevertheless, $\mathcal{K}_1$ has the correct sign and is significant, meaning that an underprediction leads to an upward revision of the trend. In case of the cyclical component of expectations, results are less clear–cut. Unfortunately, even a negative value for $\mathcal{K}_3$ occurs once. There are differences when we regard different surveys, as the model does not seem to be a good explanation of expectations contained in *MHS h=12*. There are also differences with respect to forecast horizons concerning the implied trend expectation. As far as the cyclical component is concerned, *LIV h=1* is the only case where the transient part is not explained by the model. Admittedly, the simple model is not capable to explain expectation formation perfectly, which translates into an irregular component that shows systematic movements. This is – in part – due to the fact that I employ the same model for every survey measure. In section 3.5, I will follow an approach where aggregate expectations are assumed to be heterogeneous, which is one interpretation of the result presented in the analysis above. Another implication of the present findings is, that there may be some time–variation of expectation formation schemes depending on the distinct presidential periods. One might conjecture that for the Volcker disinflation signal extraction seems to work better than for the later period. Up to now, it is still an open question how quickly agents would learn with a strict application of the Kalman filter. Consequently, the following section presents an out–of–sample simulation that allows to compare some common forecasting models with respect to forecast performance and – most important – their ability to explain survey expectations.

3.4 Forecasting Inflation

3.4.1 The Framework

In the following, I will simulate the forecasting exercise undertaken by survey participants taking an out–of–sample perspective. Equations (3.2) to (3.4) constitute the data generating process for inflation expectations, where the forecaster is assumed to behave like an econometrician. Similar to Branch and Evans (2006), the forecasting procedure can then be split up into three steps. In a first period, the forecaster gains some experience over the dynamics of inflation rates. In this in–sample period, he estimates the parameters of the model and runs the Kalman recursions to obtain estimates of the unobserved components. He also observes the updating gain implied by the Kalman filter. This in–sample estimation period starts at the beginning of 1953 and ends in 1980 for all models. Thus, the first period covers 27 years of data which should suffice to shape the experience of a forecaster – i.e. to obtain reliable estimates. In a second step, the forecaster takes the estimated parameters as given and extracts the unobserved components up to the last published record of inflation by relying on the gain parameters estimated during the first period. This is done subsequently for each observation following this in–sample period. In the third step, the forecaster then generates an out–of–sample forecast of the signal variable – i.e. inflation. The forecast horizon is chosen such that it matches the respective survey forecast.[9] Note that a number of different models have to be built, as survey expectations involve different target variables. In addition, some exogenous variables have been added to the forecasting models to account for the fact that survey participants might also look at aggregate output or interest rates when forming their forecasts. As benchmark cases, I also introduce a naive forecasting scheme (Model I) and a simple AR(1) model in first differences (Model II). The reason is, that the model does not involve trend considerations explicitly. Nevertheless, some learning takes place because parameters are estimated by recursive least squares. To be more precise, Model II comes very close to the type of learning models employed by – among others – Branch and Evans (2006). It has the features of a widely implemented (decreasing gain) learning algorithm as parameter estimates are updated every time a new observation becomes available[10]. I also introduce perfect foresight, which provides rational expectations (Model VII). In detail, the following models

[9]The forecast horizon is the next quarter (*SPF h=1*), the average of the next 4 quarters (*SPF h=4*), the next half year (*LIV h=1*), the next year (*LIV h=2*) and the average of the next 12 months (*MHS* h=12).

[10]See Evans and Honkapohja (2001) for further details on recursive least squares learning schemes. Here, it is generally assumed that private agents learn the parameter values of the rational expectations solution of the model. Also see Branch and Evans (2006) and Weber (2007) and the papers cited there for empirical approaches.

have been employed:

Model I: $E_t\pi_{t+h} = \pi_t$

Model II: $\Delta\pi_t = \alpha_0 + \alpha_1\Delta\pi_{t-1} + \varepsilon_t$ estimated recursively

Model III: $\pi_t = \overline{\pi}_t + \widehat{\pi}_t + d_t + \varepsilon_t$

Model IV: $\pi_t = \overline{\pi}_t + \widehat{\pi}_t + d_t + \sum_{\tau=0}^{3} \Delta i_{t-h-\tau} + \sum_{\tau=0}^{3} \Delta y_{t-h-\tau} + \varepsilon_t$

Model V: Model III estimated recursively

Model VI: Model IV estimated recursively

Model VII: $E_t\pi_{t+h} = \pi_{t+h}$

Here, Δi_t denotes the change of the three month treasury bill rate measured either on a quarterly frequency (*SPF*) or on a monthly frequency (*LIV*, *MHS*). Δy_t is a measure of aggregate output growth, which means that the change of industrial production has been employed on a monthly frequency and GDP growth for *SPF*, which is observed on a quarterly frequency.[11]

Models III and VI are estimated by maximum likelihood, whereas the likelihood is computed by the Kalman filter, which is initialized with a diffuse prior density of the state vector. The parameter vector, here, comprises five variables $\psi = \begin{pmatrix}\sigma_\eta^2 & \sigma_\kappa^2 & \lambda & \rho & \sigma_\varepsilon^2\end{pmatrix}$. During estimation ψ is reparameterized according to theoretical restrictions (see appendix 3.B for details). Models V and VI are essentially the same as Models II and III but estimated recursively. This means that forecasters come up with a new set of maximum likelihood estimates whenever new data becomes available. Most importantly, also the estimated gain parameters will change with every new observation. This is a more sophisticated forecasting scheme than before, as it assumes that forecasters revise their estimates from time to time. But it also represents a learning scheme where agents learn from past misperceptions.

The different forecasting models are now used to forecast five different target series, which are chosen such that they match the respective survey (see figure 3.1 for a graphical representation of the time series to be forecast). Note, that simulated forecasters use monthly data for six and twelve month CPI inflation but the forecast is made on a semi–annual frequency. However, for the twelve month average CPI inflation series, the forecast is produced every month. To ensure that the simulated forecasters start out with a well specified model in 1980, some dummy variables are introduced to the model whenever the outlier t–test proposed by Harvey and

[11] Note that π_t is the quarterly change of GDP inflation for *SPF h=1*, the average annualized GDP inflation during the next four quarters for *SPF h=4*, the annualized 6 months CPI inflation for *LIV h=1*, the twelve months CPI inflation for *LIV h=2* and the average annualized CPI inflation during the next twelve months for *MHS h=12*.

Koopman (1992) showed signs of severe outliers. In appendix 3.D, the estimated components and a couple of diagnostics is presented. On the whole, the tests give satisfying results and indicate that the forecasting exercise relies on well specified models, although, for each target variable, the same forecasting model has been used. In case of the annualized 6 month CPI and the GDP deflator, tests show some signs of a distinct structural break around 1975, not accounted for by the model.

3.4.2 Forecast Accuracy

In a next step, a test of forecast accuracy of Models I to VI is presented in table 3.3[12]. Here, negative values imply that the model in row i has a lower forecast error than the model in column j. It becomes apparent that signal extraction with the simple model performs poorly. In general, these models are dominated by simpler versions. For instance, considering the models for GDP inflation, the recursive autoregressive model (Model II) performs better than the other ones. Nevertheless, the models are very similar as the test statistic is significant only in two cases. Interestingly, estimating Models II and III recursively or adding interest rates and real output to the equations does not seem to improve forecasting ability substantially. By contrast, adding exogenous variables significantly reduces out–of–sample forecasting ability in the quarterly GDP inflation model. The picture is similar for forecasts of CPI inflation, where even the naive forecasting scheme produces the lowest forecast errors for each of the three different target variables.

On the whole, it becomes clear that signal extraction with a simple univariate model does not outperform other simple models as far as the forecasting error is concerned. This is not surprising, as in all models – in the spirit of bounded rationality – the dynamics is kept quite simple.

3.4.3 Estimated Parameters

Models V and VI are estimated recursively, whereas the resulting estimates of structural and gain parameters $\mathcal{K}_1$ to $\mathcal{K}_3$ are given in appendix 3.E in figures 3.20 to 3.23[13]. The upper part of the figure shows parameter estimates for Model V and

[12] The test has originally been proposed by Diebold and Mariano (1995) and has been augmented by Harvey, Leybourne, and Newbold (1997) to account for overlapping forecast errors and small sample bias. The null hypothesis is given by $H_0 : E[|\pi_{t+h} - \pi^{f,i}_{t+h|t}| - |\pi_{t+h} - \pi^{f,j}_{t+h|t}|] = 0$. Here, $\pi^{f,i}_{t+h|t}$ is the h–period out–of–sample forecast stemming from model i.

[13] For the presentation of gain parameters, it is generally assumed that, once having estimated the hyperparameters, the covariance matrix of innovations converges to the steady–state solution when the Kalman filter is run up to the last observation. Hence, the graphs show the estimated

	Quarterly GDP inflation annualized (h=1)						4 quarter GDP inflation average (h=4)					
Model	I	II	III	IV	V	VI	I	II	III	IV	V	VI
I	0.00	0.79	0.55	-1.18	0.50	-1.34	0.00	0.04	-0.55	-0.78	-0.46	-0.51
II	-0.79	0.00	-0.07	-1.91	-0.22	-2.23	-0.04	0.00	-0.65	-1.05	-0.57	-0.72
III	-0.55	0.07	0.00	-2.25	-0.41	-2.47	0.55	0.65	0.00	-0.64	1.04	0.14
IV	1.18	1.91	2.25	0.00	2.16	-0.14	0.78	1.05	0.64	0.00	0.82	1.14
V	-0.50	0.22	0.41	-2.16	0.00	-2.46	0.46	0.57	-1.04	-0.82	0.00	-0.08
VI	1.34	2.23	2.47	0.14	2.46	0.00	0.51	0.72	-0.14	-1.14	0.08	0.00
	6 month CPI inflation annualized (h=1)						12 month CPI inflation (h=2)					
Model	I	II	III	IV	V	VI	I	II	III	IV	V	VI
I	0.00	-2.02	-2.59	-2.06	-2.49	-2.09	0.00	-2.47	-0.86	-0.60	-0.93	-0.66
II	2.02	0.00	-1.74	-1.26	-1.65	-1.27	2.47	0.00	-0.61	-0.37	-0.69	-0.43
III	2.59	1.74	0.00	0.37	0.48	0.48	0.86	0.61	0.00	1.31	-0.45	0.93
IV	2.06	1.26	-0.37	0.00	-0.18	0.23	0.60	0.37	-1.31	0.00	-1.25	-0.88
V	2.49	1.65	-0.48	0.18	0.00	0.29	0.93	0.69	0.45	1.25	0.00	1.14
VI	2.09	1.27	-0.48	-0.23	-0.29	0.00	0.66	0.43	-0.93	0.88	-1.14	0.00
	12 month CPI inflation average (h=12)											
Model	I	II	III	IV	V	VI						
I	0.00	-1.87	-0.94	-0.93	-0.98	-0.25						
II	1.87	0.00	-0.75	-0.73	-0.79	0.02						
III	0.94	0.75	0.00	0.59	-0.05	1.12						
IV	0.93	0.73	-0.59	0.00	-0.39	1.15						
V	0.98	0.79	0.05	0.39	0.00	1.17						
VI	0.25	-0.02	-1.12	-1.15	-1.17	0.00						

Note: Numbers are modified Diebold–Mariano (DM) test statistics which follow a t–distribution with $n-1$ degrees of freedom. Here, $n = 108$ ($n = 54$ and $n = 324$) is the number of out–of–sample forecasts in the top (middle and lower) part of the table. Thus, $H_0 : DM = 0$ (equal forecast performance) can be rejected on the 5% level if the test statistic exceeds 1.98 (2.00 and 1.97) in absolute values (two–sided test). A negative number means that the model in row i has a lower measured forecast error than the model in column j of the respective panel.

Table 3.3: Modified Diebold–Mariano test on forecast properties

the lower part contains Model VI parameters. The respective left hand side panel depicts estimated variances and cycle parameters and the implied gain parameters can be observed from the right hand side graphs. On the whole, hyperparameters are stable; only estimated variances seem to display some tendency to fall over time. In the one or the other case, the simulation exercise converges to a solution that implies a jump in parameter estimates, which would then show up as a distinct peak or drop in the series. During this out–of–sample simulation, these occurrences are simply taken as given and can be interpreted as the difficulty of the respective forecaster in finding an appropriate model at each point in time. Turning now to the case of GDP inflation, gain parameters start out lower in the beginning of the estimation period with values around 0.40 and slowly rise to 0.50 when new data becomes available. However, there is some variation of $\mathcal{K}_1$ over time – in particular until around 1987. A very similar pattern emerges for the six month CPI inflation model. The twelve month CPI inflation model updates trend forecasts with a gain parameter which implies that about 70% of the error is associated with trend misperceptions. Unfortunately, if interest rates and output are added to the equations,

gain parameters conditional on the whole data set available at the time the forecast is made.

then the cycle turns out to be deterministic except during the first three years and in the beginning of the nineties.

When comparing the gain parameters which apply for the out–of–sample models and estimated learning dynamics in–sample, it becomes apparent that in an out–of–sample experiment forecasters would change their trend perceptions much more often. Consequently, this leads to trend expectations that are more volatile than those observed for the survey measures in section 3.3.3. Also note that these gain parameters are optimal within this type of model in the sense that they minimize the one–step ahead forecast error. However, one has to be careful when comparing in–sample estimates of gain parameters with the out–of–sample counterparts calculated by strict application of the Kalman filter. The reason is that – as outlined in section 3.3.3 – participants of *SPF h=4*, *LIV h=2* and *MHS h=12* are assumed to learn from multi–step forecast errors that induce an overlap which is not the case for the standard Kalman filter. Moreover, semi–annual CPI forecasts with a 6 month horizon are based on a monthly model, which allows for a trend update every month. This is not the case for the corresponding survey (*LIV h=1*) in the in–sample analysis, where a trend update is based on semi–annual observations. Consequently, only results for *SPF h=1* should be compared directly to the out–of–sample results of the present section. Here, it is apparent that agents could improve their forecast performance by putting more weight on trend shifts – i.e. increasing the gain parameter $\mathcal{K}_1$ from 0.12 to about 0.50.

3.4.4 Approximation of Survey Expectations

Having seen that forecasting properties of the proposed models is not generally better than that of simple backward–looking ones, it is now important to see whether the simulated forecast series $\pi^f_{t+h|t}$ match the survey–based measures $\pi^e_{t+h|t}$. In principal, also the approximation properties can be tested by the augmented version of the Diebold–Mariano statistic[14]. The resulting test statistics can be inferred from tables 3.4 to 3.6. In addition, the sample is split into two parts, the first one covering the whole sample 1980–2007, the middle panel covers the Volcker disinflation period 1980–1987 and the last panel covers the more moderate period 1988–2007.

[14]Here, the null hypothesis is given by $H_0 : E[|\pi^e_{t+h} - \pi^{f,i}_{t+h|t}| - |\pi^e_{t+h} - \pi^{f,j}_{t+h|t}|] = 0$.

	SPF h=1 80–07							SPF h=4 80–07						
Model	I	II	III	IV	V	VI	VII	I	II	III	IV	V	VI	VII
I	0.00	5.84	6.45	4.95	6.37	4.42	-0.61	0.00	-2.21	0.49	-1.41	0.00	-0.83	-1.98
II	-5.84	0.00	2.04	2.05	1.67	1.41	-2.44	2.21	0.00	1.51	-0.94	1.26	0.28	-1.67
III	-6.45	-2.04	0.00	1.10	-2.20	0.10	-3.16	-0.49	-1.51	0.00	-2.04	-2.13	-1.33	-2.15
IV	-4.95	-2.05	-1.10	0.00	-1.62	-2.72	-3.34	1.41	0.94	2.04	0.00	1.79	1.57	-1.09
V	-6.37	-1.67	2.20	1.62	0.00	0.64	-2.95	0.00	-1.26	2.13	-1.79	0.00	-1.01	-1.97
VI	-4.42	-1.41	-0.10	2.72	-0.64	0.00	-2.94	0.83	-0.28	1.33	-1.57	1.01	0.00	-1.78
VII	0.61	2.44	3.16	3.34	2.95	2.94	0.00	1.98	1.67	2.15	1.09	1.97	1.78	0.00
	SPF h=1 80–87							SPF h=4 80–87						
Model	I	II	III	IV	V	VI	VII	I	II	III	IV	V	VI	VII
I	0.00	3.33	3.56	2.27	3.12	1.82	-0.80	0.00	-1.85	1.15	-0.93	0.72	-0.39	-1.47
II	-3.33	0.00	1.47	0.95	0.69	0.47	-1.91	1.85	0.00	2.09	-0.61	1.72	0.40	-1.25
III	-3.56	-1.47	0.00	0.15	-4.49	-0.53	-2.68	-1.15	-2.09	0.00	-2.24	-2.78	-1.47	-2.23
IV	-2.27	-0.95	-0.15	0.00	-0.82	-2.04	-2.34	0.93	0.61	2.24	0.00	1.79	1.15	-1.00
V	-3.12	-0.69	4.49	0.82	0.00	0.15	-2.23	-0.72	-1.72	2.78	-1.79	0.00	-1.10	-1.90
VI	-1.82	-0.47	0.53	2.04	-0.15	0.00	-1.95	0.39	-0.40	1.47	-1.15	1.10	0.00	-1.52
VII	0.80	1.91	2.68	2.34	2.23	1.95	0.00	1.47	1.25	2.23	1.00	1.90	1.52	0.00
	SPF h=1 88–07							SPF h=4 88-07						
Model	I	II	III	IV	V	VI	VII	I	II	III	IV	V	VI	VII
I	0.00	4.83	6.38	6.12	6.27	5.99	-0.08	0.00	-1.43	-0.71	-1.23	-1.00	-0.93	-1.63
II	-4.83	0.00	1.59	3.43	2.53	3.36	-1.56	1.43	0.00	0.13	-0.77	-0.05	-0.17	-1.23
III	-6.38	-1.59	0.00	3.14	0.81	1.91	-1.88	0.71	-0.13	0.00	-0.99	-0.78	-0.30	-1.29
IV	-6.12	-3.43	-3.14	0.00	-2.89	-2.00	-2.37	1.23	0.77	0.99	0.00	0.89	3.19	-0.50
V	-6.27	-2.53	-0.81	2.89	0.00	1.99	-1.95	1.00	0.05	0.78	-0.89	0.00	-0.12	-1.25
VI	-5.99	-3.36	-1.91	2.00	-1.99	0.00	-2.18	0.93	0.17	0.30	-3.19	0.12	0.00	-1.07
VII	0.08	1.56	1.88	2.37	1.95	2.18	0.00	1.63	1.23	1.29	0.50	1.25	1.07	0.00

Note: Numbers are modified Diebold–Mariano (DM) test statistics which follow a t–distribution with $n-1$ degrees of freedom. Here, $n = 108$ is the number of out–of–sample forecasts for the full sample. $H_0 : DM = 0$ can be rejected on the 5% level if the test statistic exceeds 1.98 in absolute values (two–sided test). A negative number means that the model in row i has a lower measured deviation than the model in column j. The first part captures results based on the whole sample, whereas the second and third part contain results for two sub–samples split in the end of 1987.

Table 3.4: Modified Diebold–Mariano test on deviations of *SPF* and forecasting models

Model	LIV h=1 80–07 I	II	III	IV	V	VI	VII	LIV h=2 80–07 I	II	III	IV	V	VI	VII
I	0.00	0.17	0.76	0.34	1.50	0.93	-1.43	0.00	0.43	-0.30	-0.24	-0.52	-0.36	-1.20
II	-0.17	0.00	0.71	0.28	1.46	0.86	-1.39	-0.43	0.00	-0.34	-0.28	-0.55	-0.39	-1.22
III	-0.76	-0.71	0.00	-0.50	3.54	0.47	-1.52	0.30	0.34	0.00	0.42	-1.04	-0.15	-1.13
IV	-0.34	-0.28	0.50	0.00	1.38	3.36	-1.47	0.24	0.28	-0.42	0.00	-1.03	-0.61	-1.12
V	-1.50	-1.46	-3.54	-1.38	0.00	-0.54	-1.81	0.52	0.55	1.04	1.03	0.00	0.60	-0.94
VI	-0.93	-0.86	-0.47	-3.36	0.54	0.00	-1.71	0.36	0.39	0.15	0.61	-0.60	0.00	-1.07
VII	1.43	1.39	1.52	1.47	1.81	1.71	0.00	1.20	1.22	1.13	1.12	0.94	1.07	0.00
	LIV h=1 80-87							LIV h=2 80-87						
Model	I	II	III	IV	V	VI	VII	I	II	III	IV	V	VI	VII
I	0.00	0.65	1.49	1.10	1.59	1.16	-1.24	0.00	0.52	0.12	0.24	-0.09	0.04	-0.89
II	-0.65	0.00	1.31	0.90	1.42	0.97	-1.26	-0.52	0.00	0.06	0.18	-0.14	-0.01	-0.92
III	-1.49	-1.31	0.00	-0.16	0.98	-0.03	-1.72	-0.12	-0.06	0.00	0.93	-1.31	-0.37	-1.28
IV	-1.10	-0.90	0.16	0.00	0.32	0.67	-1.84	-0.24	-0.18	-0.93	0.00	-1.52	-1.57	-1.30
V	-1.59	-1.42	-0.98	-0.32	0.00	-0.21	-1.78	0.09	0.14	1.31	1.52	0.00	0.46	-1.02
VI	-1.16	-0.97	0.03	-0.67	0.21	0.00	-1.80	-0.04	0.01	0.37	1.57	-0.46	0.00	-1.12
VII	1.24	1.26	1.72	1.84	1.78	1.80	0.00	0.89	0.92	1.28	1.30	1.02	1.12	0.00
	LIV h=1 88–07							LIV h=2 88–07						
Model	I	II	III	IV	V	VI	VII	I	II	III	IV	V	VI	VII
I	0.00	-0.65	-1.24	-1.70	0.09	-0.54	-0.69	0.00	0.00	-0.98	-1.20	-0.95	-0.96	-0.90
II	0.65	0.00	-0.96	-1.38	0.38	-0.23	-0.61	0.00	0.00	-0.93	-1.17	-0.93	-0.92	-0.87
III	1.24	0.96	0.00	-1.22	3.79	1.44	-0.29	0.98	0.93	0.00	-0.77	0.13	0.48	-0.04
IV	1.70	1.38	1.22	0.00	3.35	4.01	-0.16	1.20	1.17	0.77	0.00	1.14	1.93	0.06
V	-0.09	-0.38	-3.79	-3.35	0.00	-1.25	-0.66	0.95	0.93	-0.13	-1.14	0.00	0.37	-0.06
VI	0.54	0.23	-1.44	-4.01	1.25	0.00	-0.51	0.96	0.92	-0.48	-1.93	-0.37	0.00	-0.12
VII	0.69	0.61	0.29	0.16	0.66	0.51	0.00	0.90	0.87	0.04	-0.06	0.06	0.12	0.00

Note: Numbers are modified Diebold–Mariano (DM) test statistics which follow a t–distribution with $n-1$ degrees of freedom. Here, $n = 54$ is the number of out–of–sample forecasts. $H_0 : DM = 0$ can be rejected on the 5% level if the test statistic exceeds 2.00 in absolute values (two–sided test). A negative number means that the model in row i has a lower measured deviation than the model in column j. The first part captures results based on the whole sample, whereas the second and third part contain results for two sub–samples split in the end of 1987.

Table 3.5: Modified Diebold–Mariano test on deviations of *LIV* and forecasting models

	MHS h=12 80–07						
Model	I	II	III	IV	V	VI	VII
I	0.00	-2.79	-0.99	-0.99	-0.98	0.04	-0.16
II	2.79	0.00	-0.74	-0.73	-0.73	0.44	0.21
III	0.99	0.74	0.00	0.41	0.26	1.31	0.92
IV	0.99	0.73	-0.41	0.00	-0.04	1.34	0.93
V	0.98	0.73	-0.26	0.04	0.00	1.30	0.90
VI	-0.04	-0.44	-1.31	-1.34	-1.30	0.00	-0.25
VII	0.16	-0.21	-0.92	-0.93	-0.90	0.25	0.00
	MHS h=12 80-87						
Model	I	II	III	IV	V	VI	VII
I	0.00	-2.30	-0.87	-0.80	-0.96	0.44	0.75
II	2.30	0.00	-0.70	-0.63	-0.77	0.67	0.91
III	0.87	0.70	0.00	0.95	-0.32	1.69	1.68
IV	0.80	0.63	-0.95	0.00	-0.74	1.61	1.62
V	0.96	0.77	0.32	0.74	0.00	1.99	1.88
VI	-0.44	-0.67	-1.69	-1.61	-1.99	0.00	0.61
VII	-0.75	-0.91	-1.68	-1.62	-1.88	-0.61	0.00
	MHS h=12 88–07						
Model	I	II	III	IV	V	VI	VII
I	0.00	-1.86	-0.42	-0.67	-0.16	-0.64	-1.86
II	1.86	0.00	-0.11	-0.30	0.20	-0.32	-1.66
III	0.42	0.11	0.00	-1.04	1.54	-0.32	-1.02
IV	0.67	0.30	1.04	0.00	8.12	-0.08	-0.95
V	0.16	-0.20	-1.54	-8.12	0.00	-0.74	-1.35
VI	0.64	0.32	0.32	0.08	0.74	0.00	-0.96
VII	1.86	1.66	1.02	0.95	1.35	0.96	0.00

Note: Numbers are modified Diebold–Mariano (DM) test statistics which follow a t–distribution with $n-1$ degrees of freedom. Here, $n = 324$ is the number of out–of–sample forecasts. $H_0 : DM = 0$ can be rejected on the 5% level if the test statistic exceeds 1.97 in absolute values (two–sided test). A negative number means that the model in row i has a lower measured deviation than the model in column j. The first part captures results based on the whole sample, whereas the second and third part contain results for two sub–samples split in the end of 1987.

Table 3.6: Modified Diebold–Mariano test on deviations of *MHS* and forecasting models

Turning first to table 3.4, the left part gives results for the approximation of *SPF h=1*. Considering the whole sample, Model IV clearly shows negative values for the modified Diebold–Mariano test statistic throughout and, hence, dominates the other out–of–sample forecasts. Moreover, the approximation error is even significantly lower when compared to forecasts obtained by Models I and II. This means that, for *SPF h=1*, learning by signal–extraction clearly gives a better approximation of survey expectations than a simple backward–looking forecasting scheme which is given by Model I. Moreover, it also outperforms recursive least squares learning of coefficients, which is represented by Model II. Interestingly, it also performs much better than the recursively estimated models V and VI. Also note that all learning models III to VI yield a closer approximation of *SPF h=1* than the models which are not characterized by signal extraction. As argued in section 3.2.2, rational expectation formation is a poor proxy for survey expectations. Splitting the sample does not alter the results. Considering *SPF h=4*, which has a forecasting horizon of one year, it becomes apparent that Model III – the simplest signal extraction model – yields the best approximation. The difference here is even significant with the exception of Model I. This basically remains true for the first sub–sample. However, during the moderate period after 1987 the naive model proxies *SPF h=4* closest but if tested against Models III to VI the difference is not significant. Now turning to the left part of table 3.5, it is apparent that Model V yields the approximation closest to *LIV h=1*. Again, it outperforms Models I and II significantly when the test is based on the whole sample and the first sub–sample. Survey expectations from *LIV h=1* cannot be approximated by rational expectations which perform worst of all models. Looking at the right panel, results are mostly insignificant. For the whole sample period recursive least squares learning seems to yield the smallest deviation from *LIV h=2* and, again, rational expectations perform worst. During the period of disinflation, however, signal extraction gives the best description of expectation formation as Model IV performs best in the first sub–sample. The second sub–sample confirms the results found for the whole sample period. Coming now to table 3.6, which contains results for *MHS h=12*, findings are rather mixed. During the whole period, the recursively estimated Model VI gives the closest approximation of *MHS h=12*. Thus, one could conclude that, also in this case, signal extraction provides the best explanation for survey expectations. However, results are not significantly better than those obtained from rational expectations or naive and simple autoregressive forecasting schemes. Moreover, when taking a look at the first sample period, rational expectations seem to give the best approximation for *MHS h=12*. When compared to figure 3.3 it becomes clear that, during the first period, forecast errors do not show any sign of persistence which is in contrast to the other survey measures of expectations and which may explain the last result. When looking at the second sub–sample, it is apparent that the naive forecasting scheme outperforms the other models.

To sum up, signal extraction gives a pretty good approximation of the expectation formation process. This is in particular the case for *SPF h=1* and *LIV h=1*. Here, learning by signal extraction generally outperforms other forecasting schemes. Furthermore, it gives a significantly better approximation of expectation formation than recursive least squares learning. However, it remains unclear whether, in general, expectations are better characterized by signal extraction models whose estimated parameters are updated over time. For *SPF h=1*, recursively estimated signal extraction models do not outperform learning models with fixed structural parameters, whereas for *LIV h=1* the recursively estimated model is better. Considering longer forecasting horizons of one year as in *SPF h=4* and *LIV h=2*, learning by signal extraction approximates survey expectations best at least during the Volcker period. Consequently, I conclude that agents seem to change their forecasting scheme over time, as during the second sample–period naive forecasting schemes seem to be more important. But also note, that the performance of these models is not significantly better when compared to Models III to VI. In case of *MHS h=12*, results are not that clear–cut. Insofar, the findings from section 3.3.3 are confirmed. Here again, it might play a role that this series is characterized by a large overlap of twelve periods and, hence, additional information from month to month observations should play an important role for the process of expectation formation.

3.5 Heterogeneous Expectations

Having argued that signal–extraction gives the best approximation of expectation formation processes, it will now be important to show that these learning schemes indeed give a close and valuable explanation of survey expectations. One of the findings of sections 3.3.3 and 3.4.4 is the importance of the heterogeneity of inflation expectations, as none of the models has so far been able to explain expectation formation perfectly. Consequently, each of the forecasting models of section 3.4 may play a role in aggregate expectation measures. However, it can be estimated how important the respective model is for an explanation of survey expectations. Moreover, it will be possible, with the concept of heterogeneous expectations, to test if a weighted average of different model forecasts made in section 3.4.4 matches survey expectations arbitrarily closely. The weights are chosen such that the sum of squared deviations υ_t^2 of the linear model $\pi^e_{t+h|t} = \sum_{i=I}^{VII} \beta_i \pi^{f,i}_{t+h|t} + \upsilon_t$ is minimized with respect to β_i under the restrictions $\sum_{i=I}^{VII} \beta_i = 1$ and $0 \leq \beta_i \leq 1 \; \forall \; i$. The resulting estimates are presented in table 3.7. Additionally, the explanatory power of the respective linear model for survey expectations is provided by the R^2 and a Ljung–Box Q–test for autocorrelation is also given in the last two columns. All results are presented for the same sub–samples as before.

		I	II	III	IV	V	VI	VII	R^2	$Q(1)$	$Q(4)$
SPF h=1	80–07	0.00	0.00	**0.09**	**0.65**	0.00	0.00	**0.26**	0.86	0.00	0.00
	80–87	0.00	0.00	**0.35**	**0.47**	0.00	0.00	**0.19**	0.88	0.86	0.12
	88–07	0.00	0.00	0.00	**0.71**	0.00	0.00	**0.29**	0.34	0.00	0.00
SPF h=4	80–07	**0.40**	0.00	**0.59**	0.00	0.00	0.00	**0.01**	0.87	0.00	0.00
	80–87	**0.18**	0.00	**0.82**	0.00	0.00	0.00	0.00	0.86	0.03	0.21
	88–07	**0.81**	0.00	**0.11**	**0.07**	0.00	0.00	**0.02**	0.43	0.00	0.00
LIV h=1	80–07	0.00	**0.39**	0.00	0.00	0.00	**0.27**	**0.34**	0.68	0.00	0.00
	80–87	0.00	**0.37**	0.00	**0.52**	0.00	0.00	**0.11**	0.46	0.73	0.86
	88–07	**0.52**	0.00	0.00	0.00	0.00	**0.02**	**0.46**	0.42	0.00	0.01
LIV h=2	80–07	0.00	**0.46**	**0.34**	0.00	0.00	0.00	**0.21**	0.79	0.00	0.00
	80–87	0.00	**0.47**	**0.38**	**0.07**	0.00	0.00	**0.08**	0.65	0.70	0.21
	88–07	**0.57**	0.00	0.00	0.00	0.00	0.00	**0.43**	0.56	0.00	0.00
MHS h=12	80–07	**0.19**	0.00	0.00	0.00	0.00	**0.33**	**0.48**	0.79	0.00	0.00
	80–87	**0.20**	0.00	0.00	0.00	0.00	**0.20**	**0.60**	0.85	0.00	0.00
	88–07	**0.13**	0.00	**0.28**	0.00	0.00	**0.28**	**0.30**	0.16	0.00	0.00

Note: Numbers represent weights of the respective forecast in the survey forecast. These weights are chosen such that the sum of squared deviations υ_t of the linear model $\pi^e_{t+h|t} = \sum_{i=I}^{VII} \beta_i \pi^{f,i}_{t+h|t} + \upsilon_t$ is minimized under the restrictions $\sum_{i=I}^{VII} \beta_i = 1$ and $0 \leq \beta_i \leq 1 \forall i$. The columns labeled $Q(1)$ and $Q(4)$ contain p–values for a Ljung–Box Q–test for autocorrelation up to 1 and 4 periods, respectively. The way it is conducted here, the test does not account for additional uncertainty contained in the out–of–sample forecasts $\pi^{f,i}_{t+h|t}$ which enter the model.

Table 3.7: Estimated weights

A first look at the weights for Models III to VI reveals that more than half of the participants seem to use a signal extraction type forecasting scheme. Interestingly, the results are quite robust across different surveys, although, as stated before, they comprise very different target variables and various forecasting horizons. The differences, however, occur between the two sub–samples of each survey which is in line with a time–varying behavior of respondents. The exception with respect to the estimation results is, again, *MHS* where learning plays no prominent role. Leaving *MHS* aside for the moment, it can be observed that the recursively estimated models V and VI attain zero weight in all cases except *LIV h=1*. Considering the two sub–samples separately, however, confirms the estimates of the other surveys. Results presented by Branch and Evans (2006) also point in this direction. One important result is given by the fact that signal extraction plays a prominent role in explaining survey expectations especially during the Volcker period. For *SPF*, signal extraction makes up for over 80% during the first sub–sample, whereas during the second period it attains a weight of about 70% for *SPF h=1* and only 18% for *SPF h=4*. In the case of *LIV* this tendency is even more pronounced, as fractions change from more than one half for the early period to virtually zero percent for the period after 1987. Furthermore, rational expectations get a weight which is below 50% for all surveys except *MHS*, which is in line with the results from above. Interestingly, the fraction of rational respondents is always higher for the second sub–sample. At the same time, however, also the share of naive forecasters increases for all surveys except *SPF h=1*. Note, that in this case simple backward–looking behavior represented by

Models I and II does not contribute to survey expectations. Now turning to *MHS*, results suggest that there is no prominent role for learning behavior. Agents seem to be either naive forecasters or rational. Only one third is found to use Model VI for forecasting.

Looking at the last three columns of table 3.7, it becomes apparent that the explanatory power of the estimated linear relationships is quite high. The R^2 suggests that more than 80% of the variation in case of *SPF* and about 70% of the variation in case of *LIV* can be explained. Taking sub–samples into account yields another interesting result. The explanatory power of the estimated relationships is higher for the early disinflation period compared to the years after 1987. Moreover, there are no signs of autocorrelation in the estimations covering the Volcker period[15]. Although R^2 is also fairly high for *MHS* during the first period, the residual is still autocorrelated. Dynamics of the second period are also not well described by heterogeneous expectations. This finding basically confirms the results presented in tables 3.2 and 3.6.

On the whole, the concept of heterogeneous expectations with a prominent role for signal extraction is well suited to explain survey measures of inflation expectations. During phases of disinflation, the model has more explanatory power than in tranquil periods, as, during the Volcker period, the R^2 is higher, the unexplained part is free of autocorrelation and the role of signal extraction is even more prominent.

3.6 Conclusions

In a first step, I have shown that the behavior of surveys on inflation expectations is not compatible with the concept of rational expectations. Survey expectations are characterized by temporary bias and considerable persistence of forecast errors. Many theoretical studies emphasize the importance of persistence of inflation expectations for the dynamics of the inflation rate. Moreover, theoretical models that assume rational expectations unrealistically predict a jump of inflation expectations following a change of the inflation target. Most importantly, such a behavior of inflation expectations cannot explain why disinflation is costly in a purely forward–looking framework. As far as the behavior of private agents is concerned, it is also

[15] The way it is conducted here, the test does not account for additional uncertainty contained in the out–of–sample forecasts $\pi^{f,i}_{t+h|t}$ which enter the model as explanatory variables. Therefore, standard parameter distributions and test statistics do not apply in this case. However, I use standard autocorrelation tests to test for systematic behavior of υ_t. This can be justified by the fact that the test is constructed with a null hypothesis of no autocorrelation and, thus, will reject too often if additional estimation uncertainty is not taken account of.

important to note that they are not assumed to be completely ignorant. By contrast, they are confronted with a difficult forecasting problem. The reason is that the inflation target pursued by the central bank is not directly observable but has to be estimated from a noisy signal.

One possible solution to this signal extraction problem is given by the Kalman filtering framework which constitutes the learning rule of private agents. To be more precise, I assume that agents estimate the *trend plus cycle model* proposed by Harvey (1989) to infer trend shifts and transitory movements. It can be shown that it is possible to fit such a model to inflation expectations of *SPF* and *LIV*. The in–sample results suggest rather slow learning of trends which can explain the sluggishness of U.S. inflation expectations.

In a next step, I conduct an out–of–sample forecasting exercise to simulate a forecaster that solves the signal extraction problem by Kalman filtering. In detail, I employ seven different models or type of forecasters which comprise the naive forecaster, learning by recursive least squares, different types of learning by Kalman filtering and a rational forecaster. It turns out that learning by Kalman filtering approximates U.S. survey expectations closest – at least during the presidency of Volcker. This holds true for several surveys comprising several target variables. Finally, in the spirit of heterogeneous expectations, I construct a weighted average of the employed forecasting schemes. It turns out that the concept of heterogeneous expectations with a prominent role for signal extraction is well suited to explain survey measures of inflation expectations. Moreover, there seems to be a change of forecasting schemes over time as the model provides a better fit during the Volcker period. The R^2 is higher, the unexplained part is free of autocorrelation and the role of signal extraction is even more prominent.

On the whole, learning in an uncertain environment provides a good explanation for the sluggishness of inflation expectations. Moreover, a large fraction of agents seems to solve some signal extraction problem during phases of disinflation. However, it will be worthwile to look at other expectation measures such as market based expectations observed in financial markets. Naturally, the use of individual data should also provide additional insight. Moreover, it will be interesting to do the in–sample analysis in a multivariate context where inflation expectations emerge from some type of Phillips curve. Of course, also out–of–sample forecasts can be generated by a multivariate model that uses information from other macroeconomic variables.

Appendix

3.A Reparameterization of Variables in Section 3.3.3

The parameters contained in ψ are reparameterized such that they obey the theoretical restrictions. The parameter vector estimated by maximum likelihood is denoted by θ. It is transformed by a vector of functions $\mathbf{g}(\theta)$ in the following way:

$$\psi = \begin{pmatrix} \lambda \\ \rho \\ \sigma_\varepsilon^2 \end{pmatrix} \equiv \mathbf{g}(\theta) = \begin{pmatrix} g_1(\theta) \\ g_2(\theta) \\ g_3(\theta) \end{pmatrix} = \begin{pmatrix} \exp(\theta_1) \\ \Phi(\theta_2) \\ \exp(2\theta_3) \end{pmatrix}. \tag{3.9}$$

In the second step we need to calculate the standard errors of the estimates. It can be shown that the transformed estimates are asymptotically normal with estimated variance $\widehat{\text{var}}(\widehat{\psi}) = G_{\widehat{\theta}}\, \widehat{\text{var}}(\widehat{\theta})\, G'_{\widehat{\theta}}$.[16] This yields the following adjustment matrix of first derivatives $G_{\widehat{\theta}}$[17].

$$\begin{aligned} G_{\widehat{\theta}} &= \begin{pmatrix} \frac{\partial g_1(\widehat{\theta})}{\partial \theta_1} & \cdots & \frac{\partial g_1(\widehat{\theta})}{\partial \theta_3} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_3(\widehat{\theta})}{\partial \theta_1} & \cdots & \frac{\partial g_3(\widehat{\theta})}{\partial \theta_3} \end{pmatrix} \\ &= \begin{pmatrix} \exp(\widehat{\theta}_1) & 0 & 0 \\ 0 & \phi(\widehat{\theta}_2) & 0 \\ 0 & 0 & 2\exp(2\widehat{\theta}_3) \end{pmatrix}. \end{aligned}$$

3.B Reparameterization of Variables in Section 3.4

The parameters contained in ψ are reparameterized such that they obey the theoretical restrictions. The parameter vector estimated by maximum likelihood is denoted

[16] See for example Kim and Nelson (1999), chapter 2.

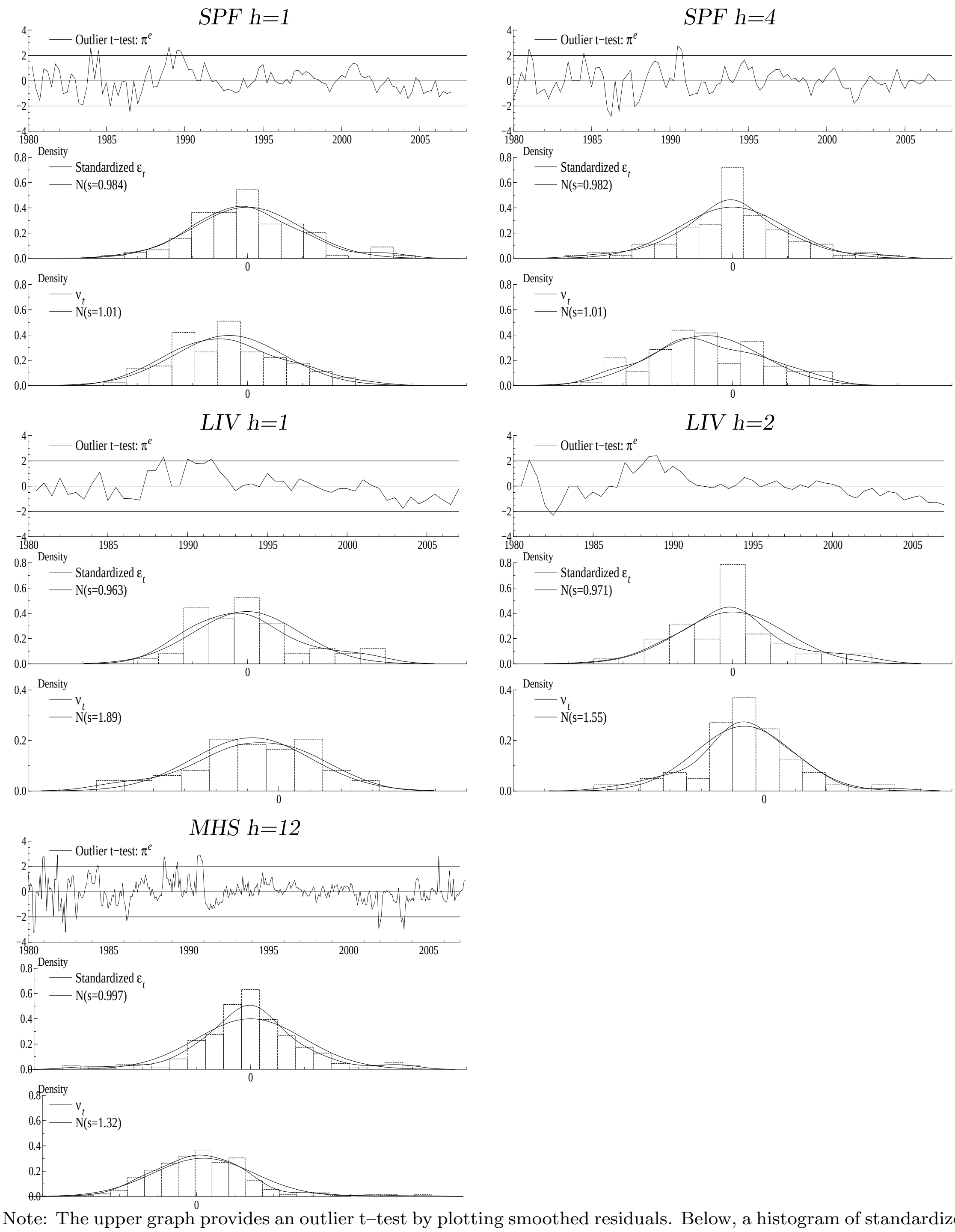
[17] The calculation of the standard error of the transformed estimates was done with the *Delta Method* which relies on first–order Taylor expansions of non–linear functions. For an overview compare Davidson and MacKinnon (2004) chapter 5.6.

by θ. It is transformed by a vector of functions $\mathbf{g}(\theta)$ in the following way:

$$\psi = \begin{pmatrix} \sigma^2_{\eta,2} \\ \sigma^2_{\kappa} \\ \lambda \\ \rho \\ \sigma^2_{\varepsilon} \end{pmatrix} \equiv \mathbf{g}(\theta) = \begin{pmatrix} g_1(\theta) \\ g_2(\theta) \\ g_3(\theta) \\ g_4(\theta) \\ g_5(\theta) \end{pmatrix} = \begin{pmatrix} \exp(2\theta_1) \\ \exp(2\theta_2)(1-\Phi(\theta_4)^2) \\ \exp(\theta_3) \\ \Phi(\theta_4) \\ \exp(2\theta_5) \end{pmatrix}. \tag{3.10}$$

3.C Diagnostics of the Learning Model

Figure 3.9 shows some diagnostics for the learning model. It depicts the respective standardized irregular component along with the associated histogram. The respective third panel shows a histogram of forecast errors for comparison.



Note: The upper graph provides an outlier t–test by plotting smoothed residuals. Below, a histogram of standardized errors ε_t can be found. The last panel of the respective graph plots the distribution of observed forecast errors ν_t.

Figure 3.9: Diagnostics learning model

3.D Diagnostics of the Forecasting Model

The following figures 3.10 to 3.18 contain smoothed unobserved components as estimated by the forecasting models and diagnostics. The upper part of the graph shows the estimated trend component along with the original series. The estimated cyclical component can be found below. The last panel graphs the irregular component. The lower panel depicts an outlier t–test, a break test which indicates distinct breaks in the mean of the series not covered by the model. Furthermore, histograms of the three standardized residuals in the system are presented, as well as the empirical autocorrelation of the innovations obtained from the Kalman filtering recursions.

Annualized quarterly GDP inflation

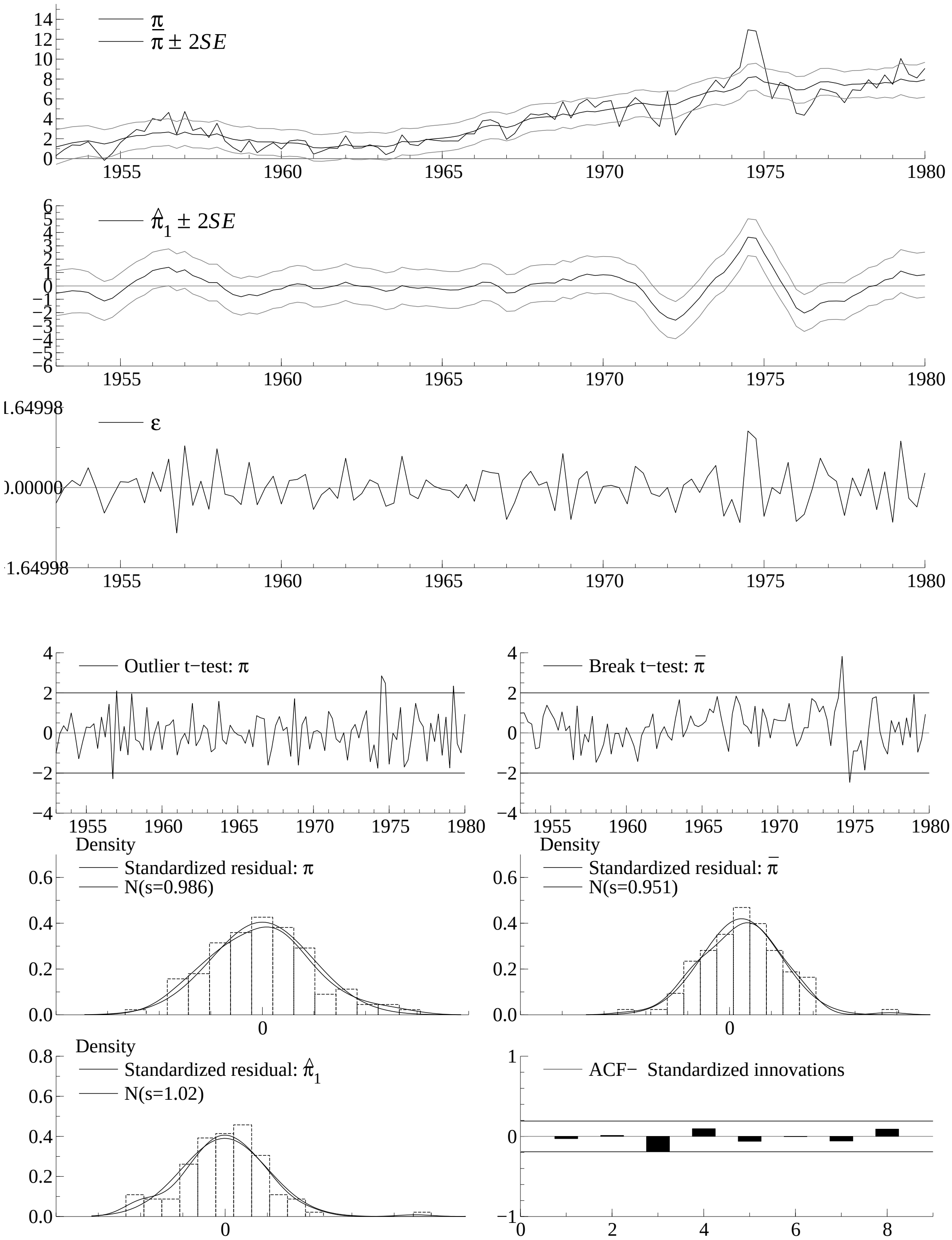


Figure 3.10: Diagnostics Model III (1)

Annualized quarterly GDP inflation

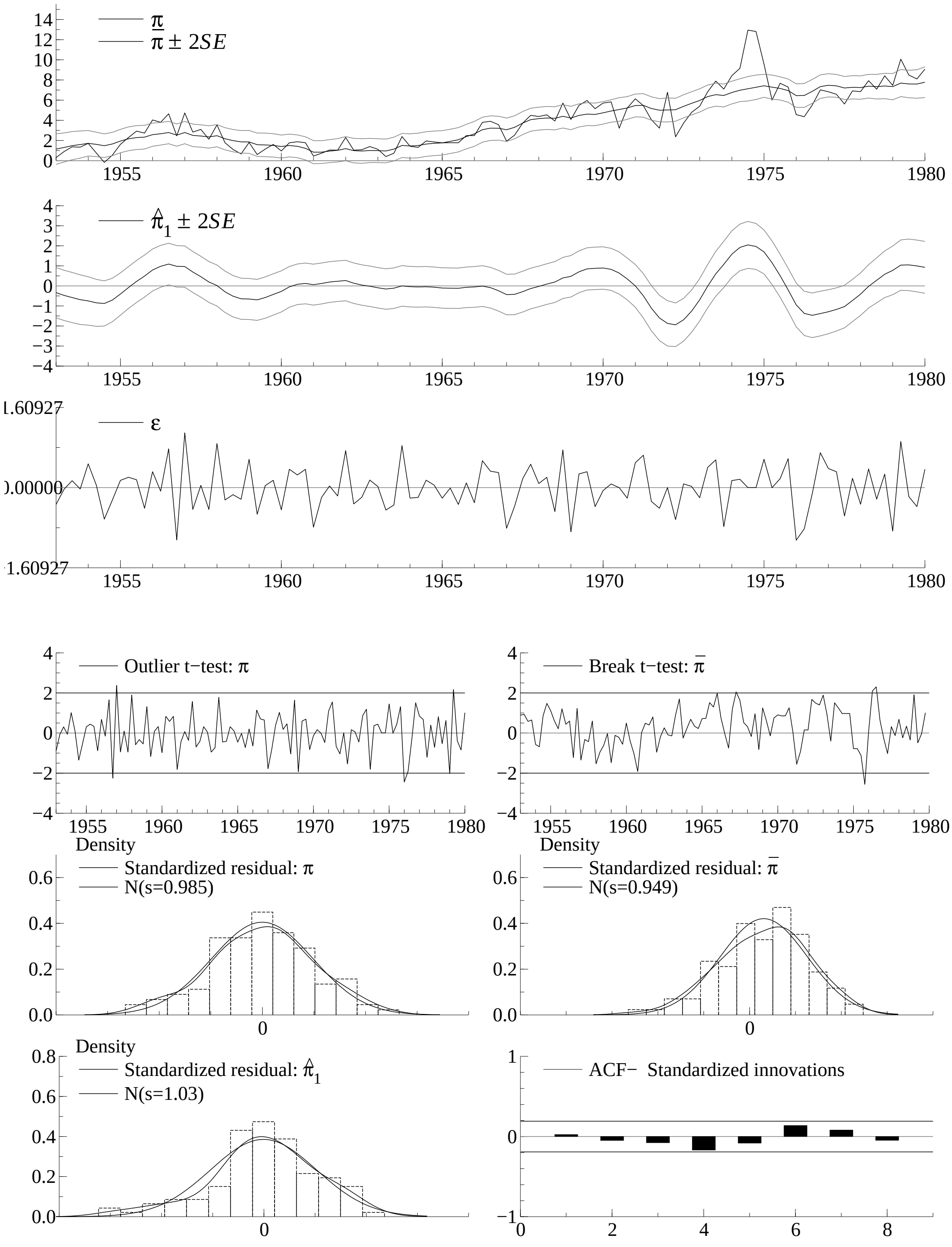


Figure 3.11: Diagnostics Model IV (1)

Average annualized 4 quarter GDP inflation

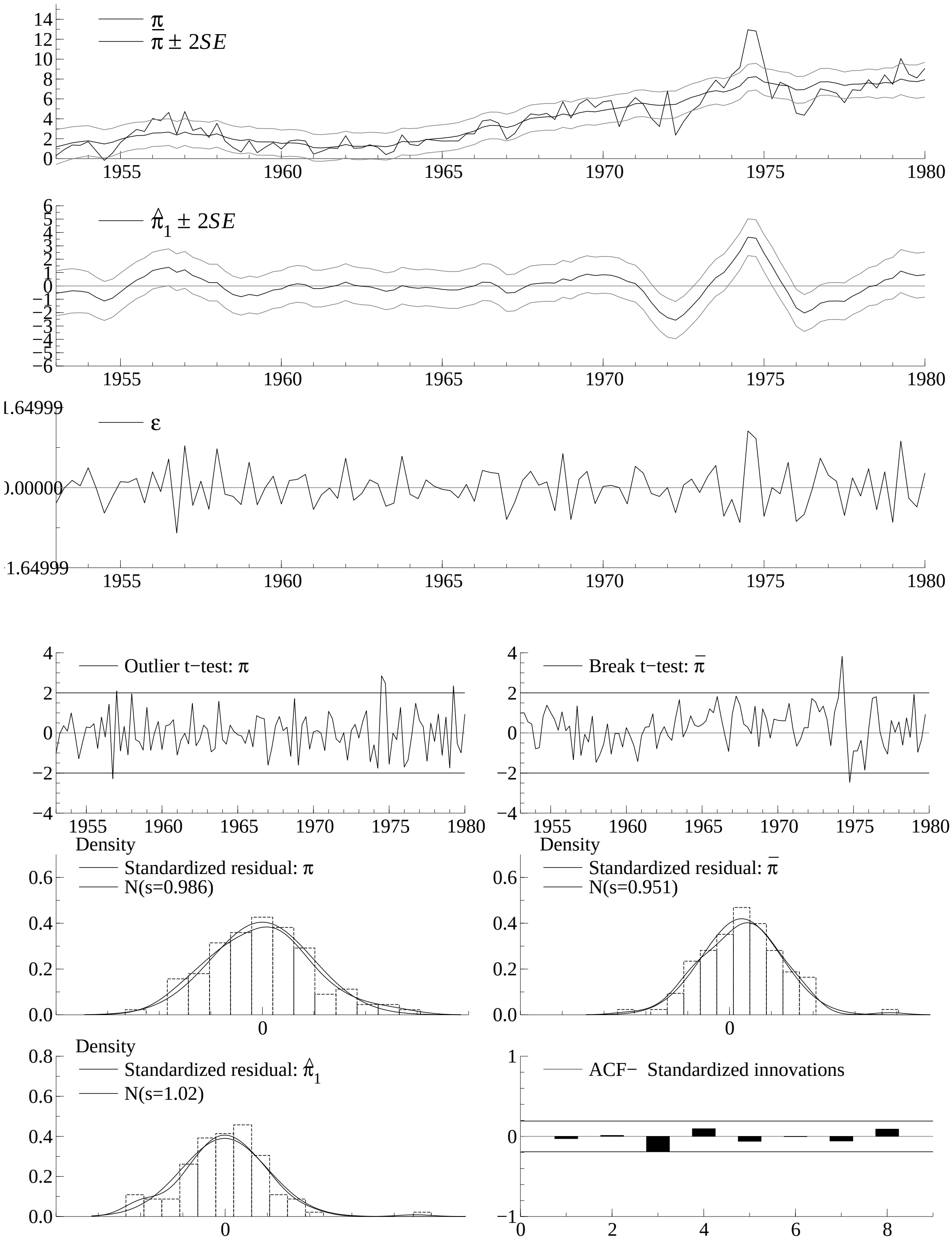


Figure 3.12: Diagnostics Model III (2)

Average annualized 4 quarter GDP inflation

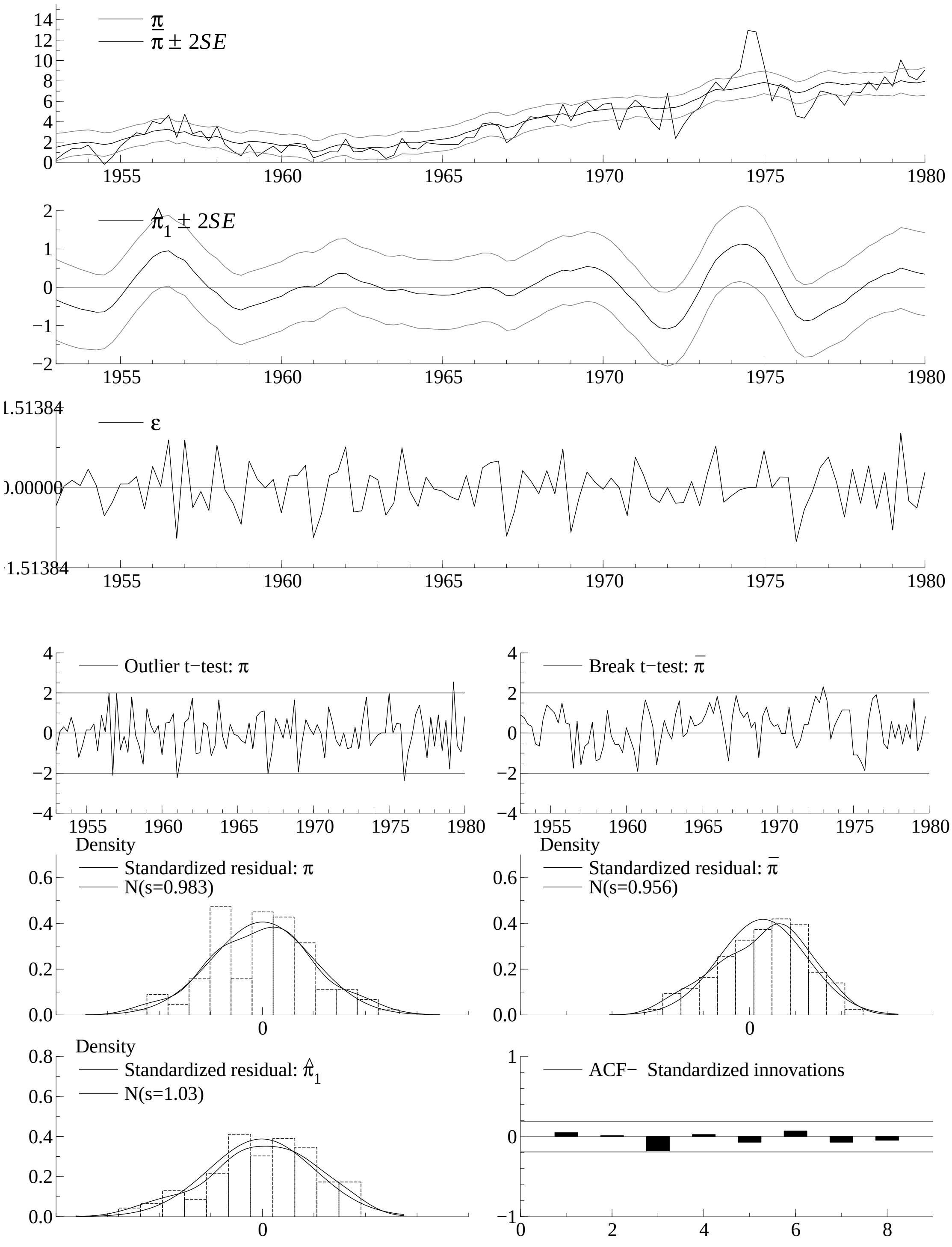


Figure 3.13: Diagnostics Model IV (2)

Annualized 6 month CPI inflation

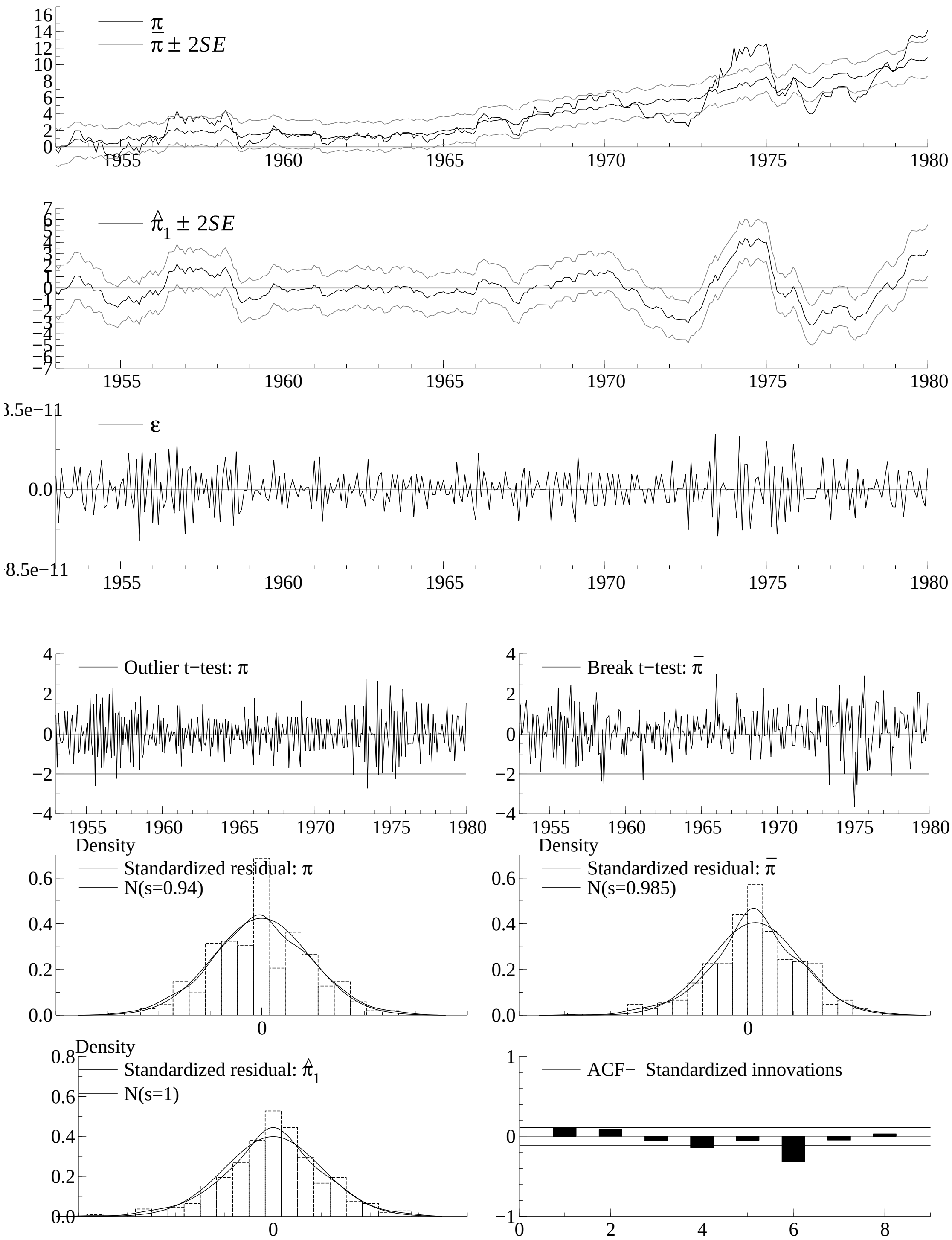


Figure 3.14: Diagnostics Model III (3)

Annualized 6 month CPI inflation

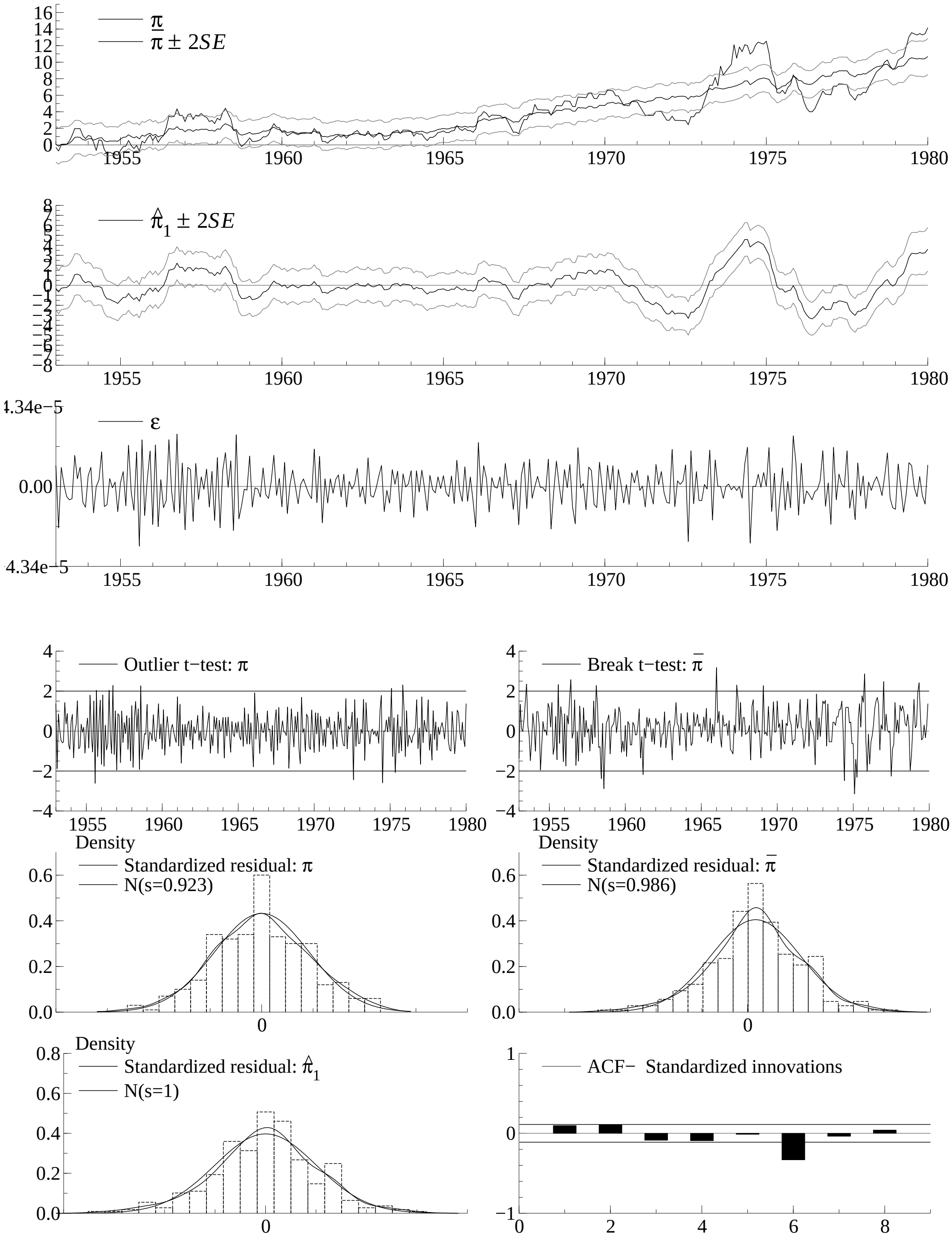


Figure 3.15: Diagnostics Model IV (3)

12 month CPI inflation

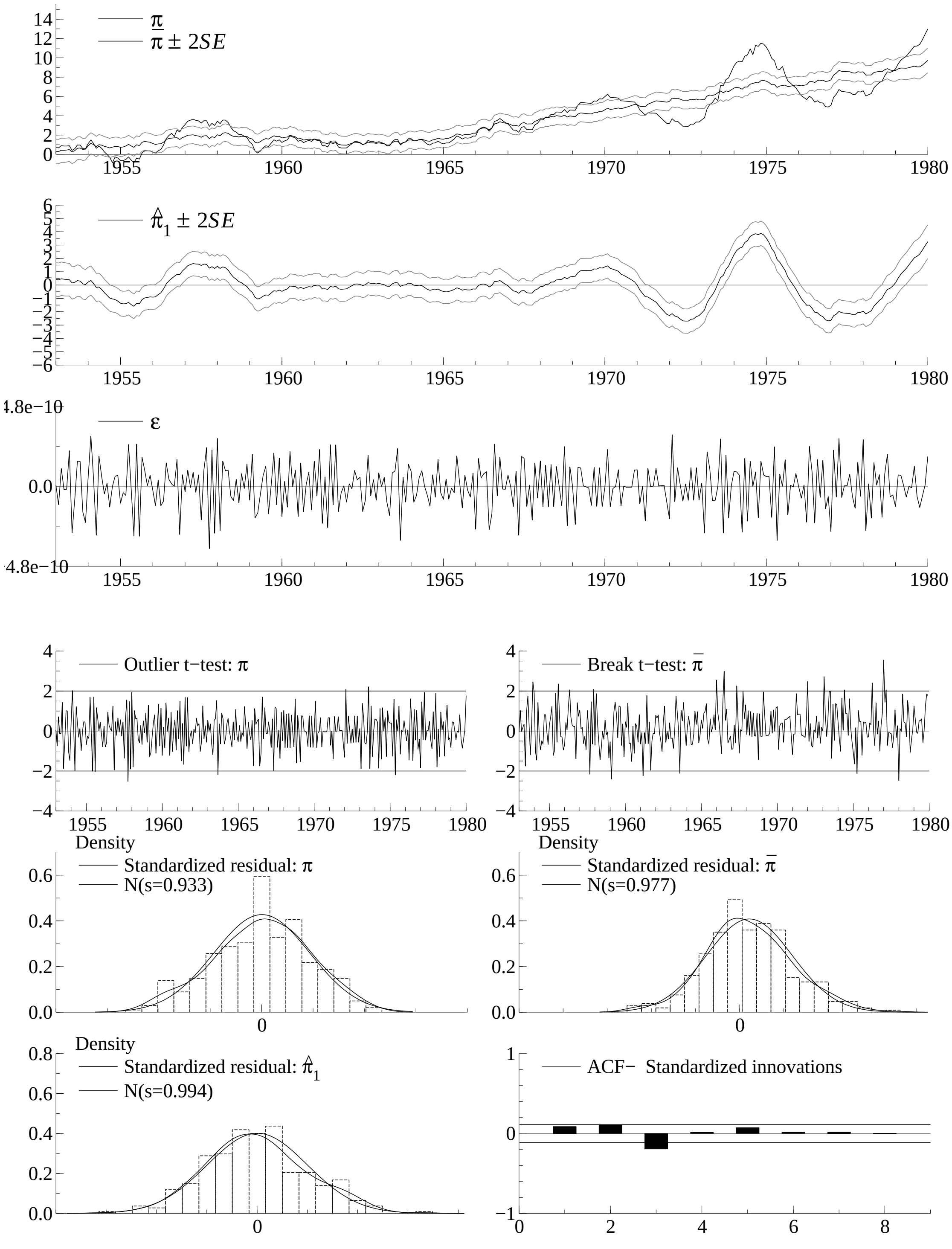


Figure 3.16: Diagnostics Model III (4)

12 month CPI inflation

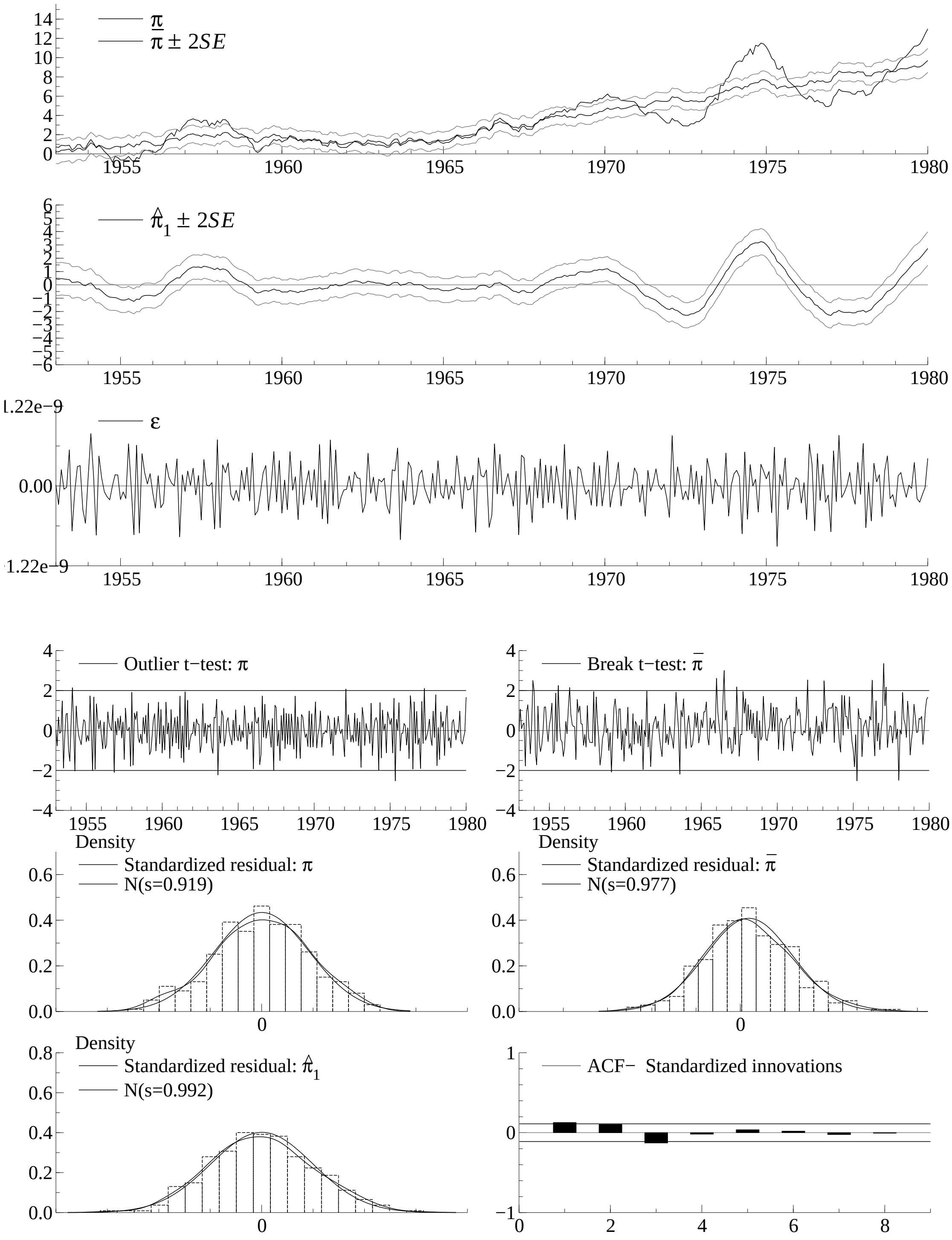


Figure 3.17: Diagnostics Model IV (4)

12 month CPI inflation

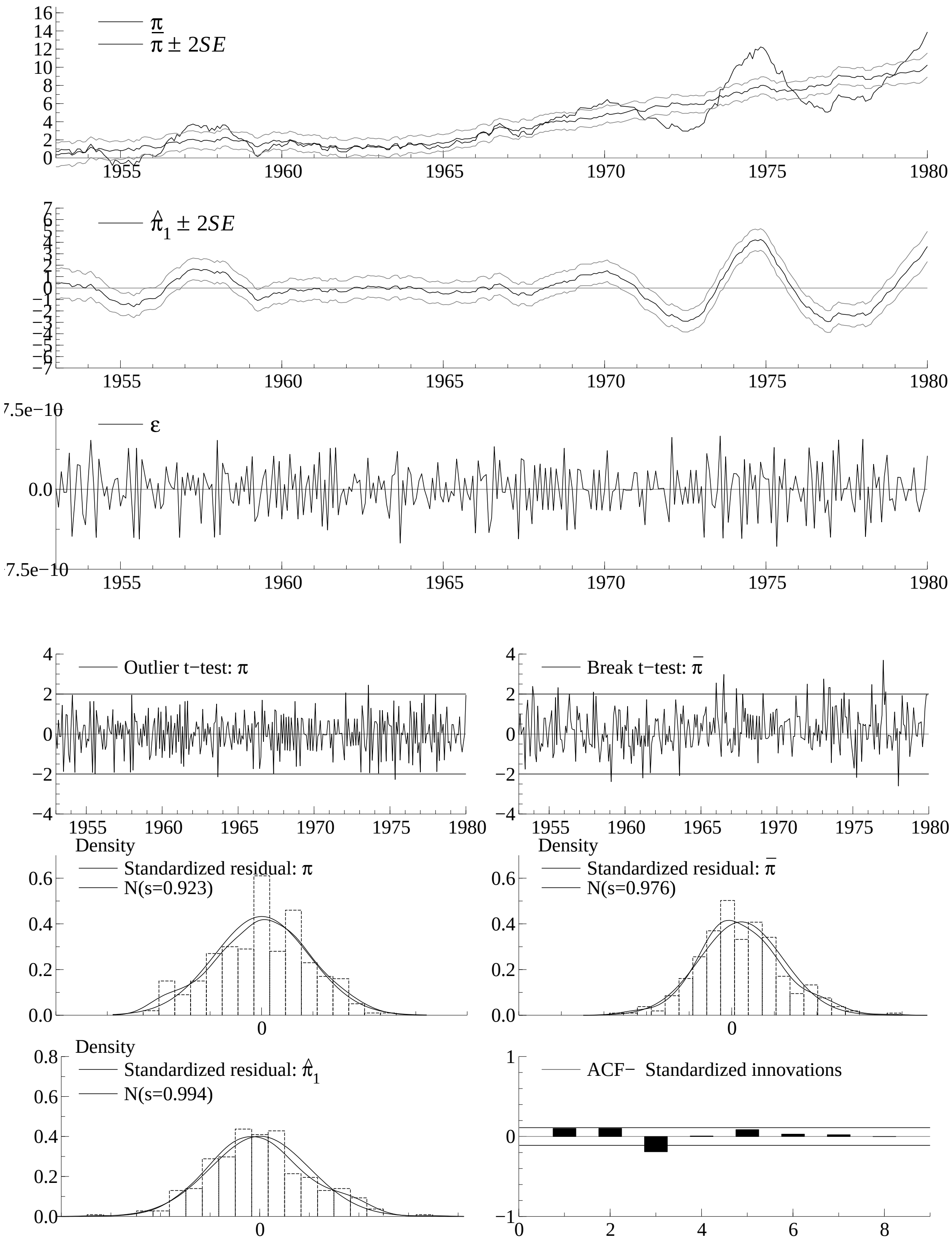


Figure 3.18: Diagnostics Model III (5)

12 month CPI inflation

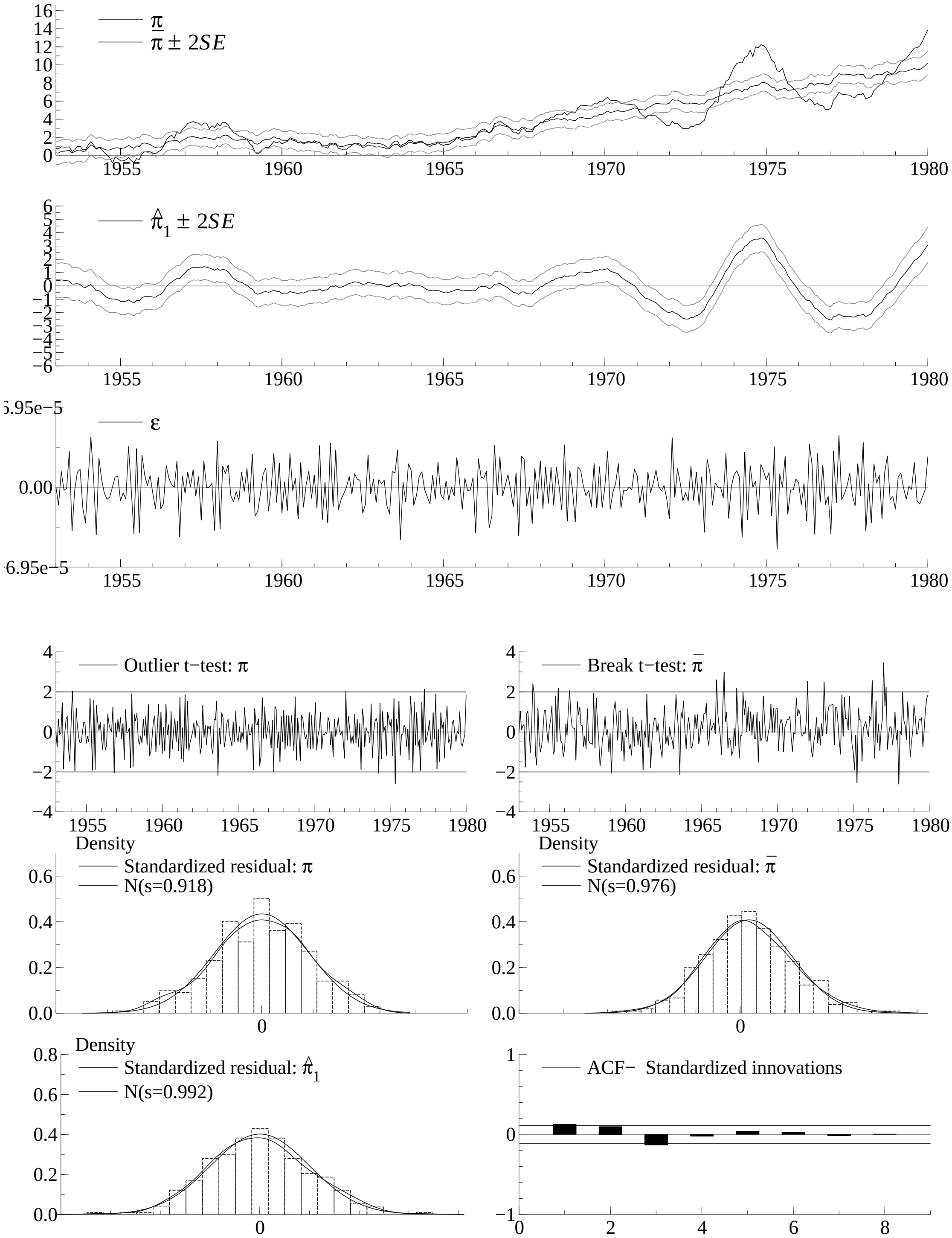


Figure 3.19: Diagnostics Model IV (5)

3.E Recursive Parameter Estimates

Figures 3.20 to 3.24 depict structural parameter estimates (left panel) along with steady–state gain parameters taken from the state vector of the system described by equations (3.2) to (3.4) (right panel). The upper part of the respective graph shows estimates from Model V whereas estimates for Model VI are presented in the lower panel.

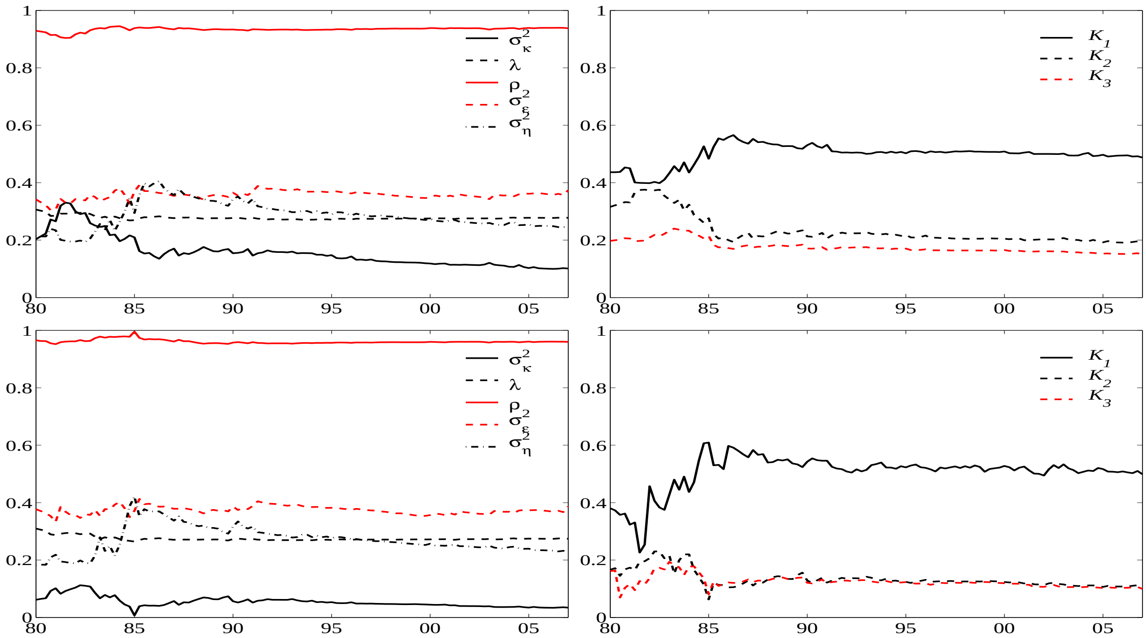


Figure 3.20: Recursively estimated parameters, *annualized quarterly GDP inflation h=1*

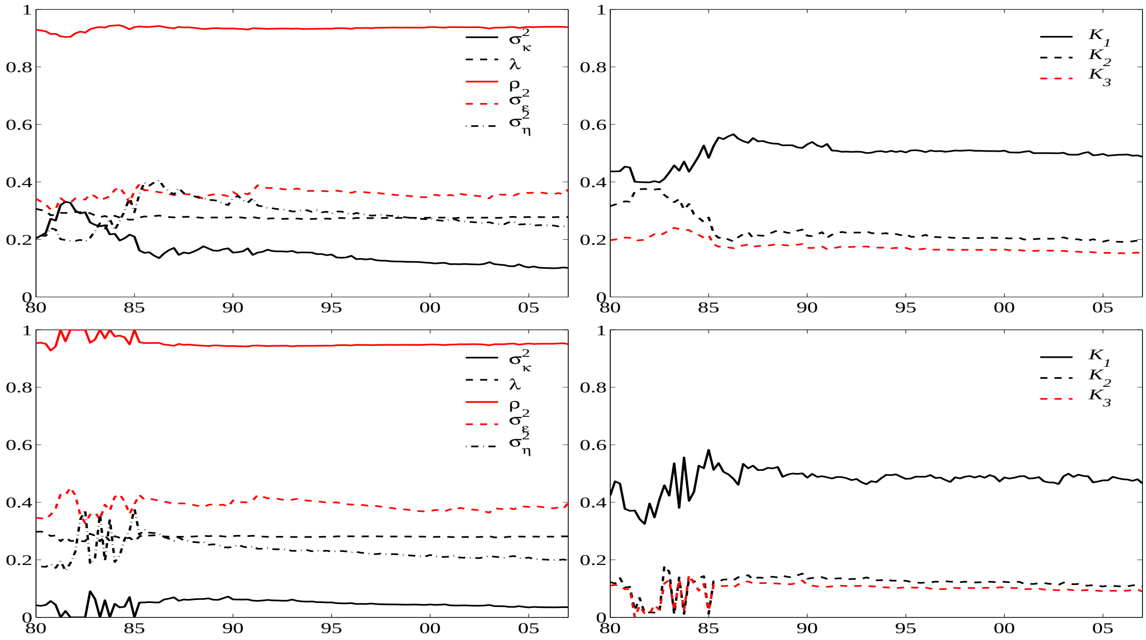


Figure 3.21: Recursively estimated parameters, *average annualized 4 quarter GDP inflation h=4*

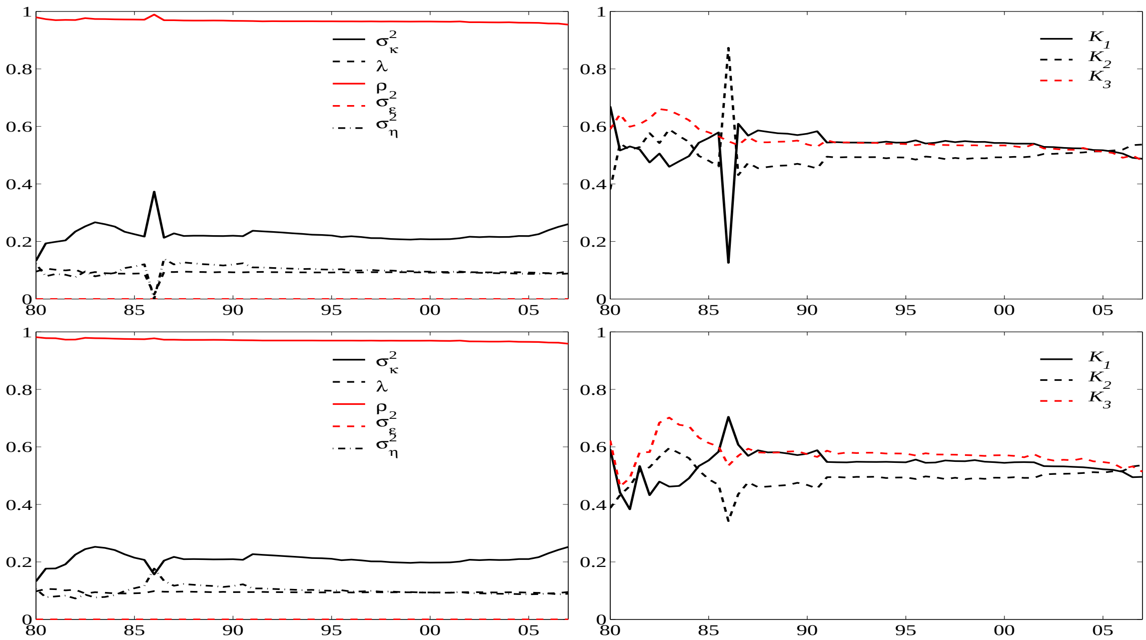


Figure 3.22: Recursively estimated parameters, *annualized 6 month CPI inflation h=1*

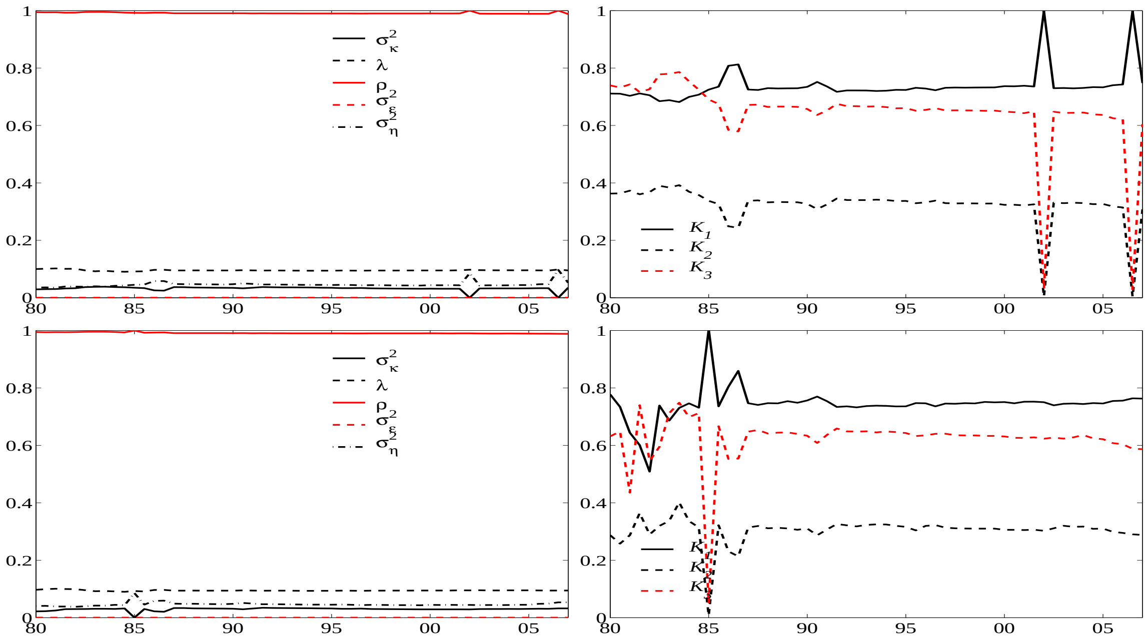


Figure 3.23: Recursively estimated parameters, *12 month CPI inflation h=2*

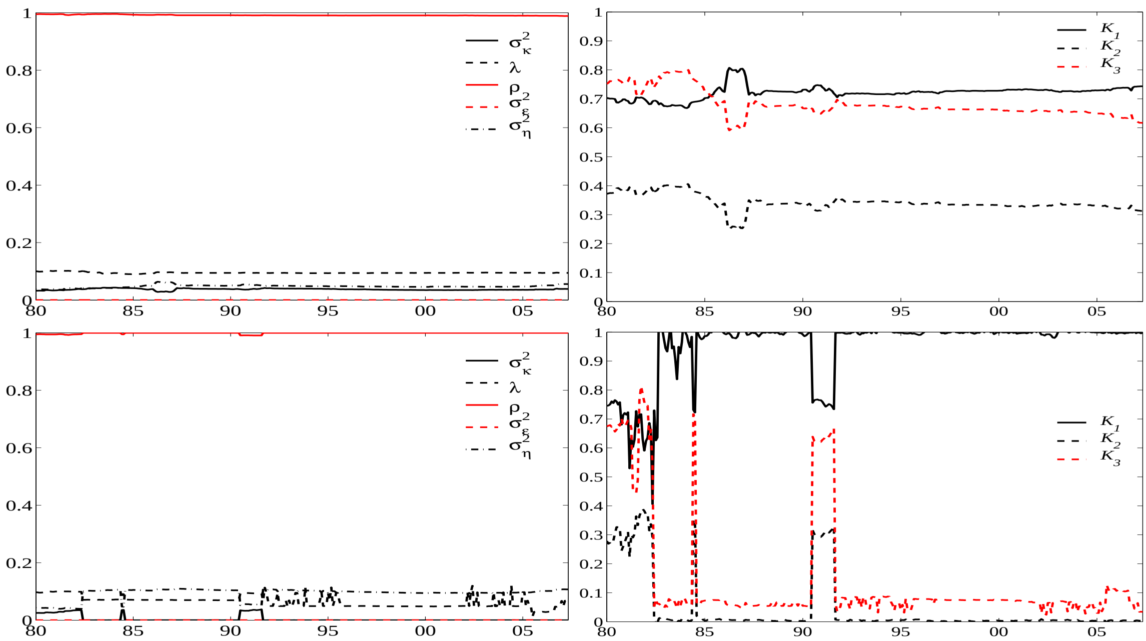


Figure 3.24: Recursively estimated parameters, *12 month average CPI inflation h=12*

References

Adam, K., and M. Padula (2003): “Inflation dynamics and subjective expectations in the United States,” ECB Working Paper No. 222.

Andolfatto, D., and P. Gomme (2003): “Monetary policy regimes and beliefs,” *International Economic Review*, 44(1), 1–30.

Andolfatto, D., S. Hendry, and K. Moran (2002): “Inflation expectations and learning about monetary policy,” *Bank of Canada Working Paper*, 30.

——— (2006): “Are inflation expectations rational?,” manuscript.

Angeloni, I., and M. Ehrmann (2004): “Euro area inflation differentials,” *The B. E. Journal of Macroeconomics*, 7, Article 24.

Ball, L. (1995): “What determines the sacrifice ratio?,” in *Monetary Policy*, ed. by N. Mankiw, pp. 155–182. Chicago University Press, Chicago.

Ball, L., N. G. Mankiw, and R. Reis (2005): “Monetary policy for inattentive economies,” *Journal of Monetary Economics*, 52(4), 703–725.

Batchelor, R. A. (1986): “The psychophysics of inflation,” *Journal of Economic Psychology*, 7, 269–290.

Batchelor, R. A., and A. B. Orr (1988): “Inflation expectations revisited,” *Economica*, 55(219), 317–331.

Benigno, P., and J. D. López-Salido (2006): “Inflation persistence and optimal monetary policy in the Euro area,” *Journal of Money, Credit, and Banking*, 38(3), 587–614.

Berk, J. M. (1999): “Measuring Inflation Expectations: A Survey Data Approach,” *Applied Economics*, 31(11), 1467–1480.

Berk, J. M. (2000): “Consumers’ inflation expectations and monetary policy in Europe,” DNB Staff Reports No. 55.

Branch, W. A. (2004): "The theory of rationally heterogeneous expectations: Evidence from survey data on inflation expectations," *The Economic Journal*, (114), 592–621.

Branch, W. A., and G. W. Evans (2006): "A simple recursive forecasting model," *Economics Letters*, 91, 158–166.

Calvo, G. A. (1983): "Staggered contracts in a utility-maximizing framework," *Journal of Monetary Economics*, 12(3), 383–398.

Carlson, J., and M. Parkin (1975): "Inflation expectations," *Economica*, 42(166), 123–138.

Cogley, T., and A. M. Sbordone (2006): "Trend inflation and inflation persistence in the New Keynesian Phillips curve," *Federal Reserve Bank of New York Staff Reports*, (270).

Cragg, J. G., and S. G. Donald (1993): "Testing identifiability and specification in instrumental variable models," *Econometric Theory*, 9, 222–240.

Cukierman, A., and A. H. Meltzer (1986): "A theory of ambiguity, credibility, and inflation under discretion and asymmetric information," *Econometrica*, 54(5), 1099–1128.

Cumby, R. E., and J. Huizinga (1992): "Testing the autocorrelation structure of disturbances in ordinary least squares and instrumental variables regressions," *Econometrica*, 60(1), 185–195.

Davidson, R., and J. G. MacKinnon (2004): *Econometric theory and methods.* Oxford University Press.

de Walque, G., F. Smets, and R. Wouters (2006): "Price shocks in general equilibrium: Alternative specifications," *CESifo Economic Studies*, 52(1), 153–176.

Diebold, F. X., and R. S. Mariano (1995): "Comparing predictive accuracy," *Journal of Business & Economic Statistics*, 13(3), 253–263.

Dossche, M., and G. Everaert (2005): "Measuring inflation persistence – A structural time series approach," ECB Working Paper No. 495.

Entorf, H. (1990): *Multisektorale Konjunkturanalyse - Theorie, Empirie und Frühindikatoren realer Konjunkturzyklen.* Campus - Verlag.

Erceg, C. J., and A. T. Levin (2003): "Imperfect credibility and inflation persistence," *Journal of Monetary Economics*, 50, 915–944.

Estrella, A., and J. C. Fuhrer (2002): "Dynamic inconsistencies: Counterfactual implications of a class of rational–expectations models," *The American Economic Review*, 92(4), 1013–1028.

Evans, G. W., and S. Honkapohja (2001): *Learning and expectations in macroeconomics*. Princeton University Press.

——— (2003): "Adaptive learning and monetary policy design," *Journal of Money, Credit and Banking*, 35(6), 1045–1072.

Evans, M., and P. Wachtel (1992): "Inflation regimes and the sources of inflation uncertainty," *Journal of Money, Credit and Banking*, 25(3), 475–511.

Fechner, G. T. (1889): *Elemente der Psychophysik*, vol. 1. Leipzig: Druck und Verlag von Breitkopf und Härtel.

Fishe, R. P., and K. Lahiri (1981): "On the estimation of inflationary expectations from qualitative responses," *Journal of Econometrics*, 16(16), 89–102.

Forsells, M., and G. Kenny (2004): "The rationality of consumers' inflation expectations: Survey-based evidence for the Euro area," *Journal of Business Cycle Measurement and Analysis*, 1(1), 13–41.

Fuhrer, J. C., and G. Moore (1995): "Inflation persistence," *The Quarterly Journal of Economics*, 110(1), 127–159.

Fuhrer, J. C., G. R. Moore, and S. D. Schuh (1995): "Estimating the linear-quadratic inventory model: Maximum likelihood versus generalized method of moments," *Journal of Monetary Economics*, 35(1), 115–157.

Galí, J. (2003): "New perspectives on monetary policy, inflation, and the business cycle," in *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress, Volume III*, ed. by M. Dewatripont, L. P. Hansen, and S. J. Turnovsky, pp. 151–197. Cambridge University Press , Cambridge.

Galí, J., and M. Gertler (1999): "Inflation dynamics: A structural econometric analysis," *Journal of Monetary Economics*, 44(2), 195–222.

Galí, J., M. Gertler, and J. D. López-Salido (2001): "European inflation dynamics," *European Economic Review*, 44(7), 1237–1270.

Giorno, C., P. Richardson, D. Roseveare, and P. van den Noord (1995): "Estimating potential output, output gaps and structural budget balances," OECD Economics Department Working Paper No. 152.

Hamilton, J. D. (1994): *Time series analysis.* Princeton University Press.

Harvey, A. C. (1989): *Forecasting, structural time series models and the Kalman filter.* Cambridge University Press.

Harvey, A. C., and S. J. Koopman (1992): "Diagnostic checking of unobserved components time series models," *Journal of Business & Economic Statistics*, 10(4), 377–389.

Harvey, D. I., S. J. Leybourne, and P. Newbold (1997): "Testing the equality of prediction mean squared errors," *International Journal of Forecasting*, 13, 281–291.

Hausman, J. A. (1978): "Specification tests in econometrics," *Econometrica*, 46(6), 1251–1271.

Henzel, S. (2008): "Learning trend inflation – Can signal extraction explain survey forecasts?," ifo Working Paper 55.

Henzel, S., and T. Wollmershäuser (2005): "Quantifying inflation expectations with the Carlson–Parkin method: A survey–based determination of the just noticeable difference," *Journal of Business Cycle Measurement and Analysis*, 2(3), 321–352.

——— (2008): "The New Keynesian Phillips curve and the role of expectations: Evidence from the CESifo World Economic Survey," *Economic Modelling*, forthcoming.

Jondeau, E., and H. LeBihan (2005): "Testing for the New Keynesian Phillips curve: Additional international evidence," *Economic Modelling*, 22(3), 521–550.

Kim, C.-J., and C. R. Nelson (1999): *State–space models with regime switching.* MIT Press, Cambridge.

Kozicki, S., and P. A. Tinsley (2005): "Permanent and transitory policy shocks in an empirical macro model with asymmetric information," *Journal of Economic Dynamics & Control*, 29(11), 1985–2015.

Leith, C., and J. Malley (2007): "Estimated open economy New Keynesian Phillips curves for the G7," *Open Economies Review*, 18(4), 405–426.

Mankiw, N. G., and R. Reis (2002): "Sticky information versus sticky prices: A proposal to replace the New Keynesian Phillips curve," *Quarterly Journal of Economics*, 117(4), 1295–1328.

MAVROEIDIS, S. (2005): "Identification issues in forward–looking models estimated by GMM, with an application to the Phillips curve," *Journal of Money, Credit, and Banking*, 37(3), 421–448.

MCCALLUM, B. T. (1976): "Rational expectations and the estimation of econometric models: An alternative procedure," *International Economic Review*, 17(2), 484–490.

MUTH, J. F. (1961): "Rational expectations and the theory of price movements," *Econometrica*, 29(3), 315–335.

NARDO, M. (2003): "The quantification of qualitative survey data: A critical assessment," *Journal of Economic Surveys*, 17(5), 645–668.

NUNES, R. C. (2004): "Learning the inflation target," manuscript, Universitat Pompeu Fabra.

——— (2005): "Inflation dynamics: The role of expectations," manuscript, Universitat Pompeu Fabra.

PAJFAR, D., AND E. SANTORO (2007): "Heterogeneity and learning in inflation expectation formation: An empirical assessment," manuscript, University of Cambridge.

PALOVIITA, M. (2006): "Inflation dynamics in the Euro area and the role of expectations: Further results," *Empirical Economics*, 31(4), 847–860.

PESARAN, M. H. (1984): "Expectations formations and macroeconomic modelling," in *Contemporary Macroeconomic Modelling*, ed. by P. Malgrange, and P. Muet, pp. 27–55. Blackwell, Oxford.

——— (1987): *The limits to rational expectations*. Basil Blackwell, Oxford.

PESARAN, M. H., AND M. WEALE (2006): "Survey expectations," in *Handbook of Economic Forecasting*, ed. by G. Elliott, C. Granger, and A. Timmermann. North-Holland, Amsterdam.

RECKWERTH, J. (1997): "Der Zusammenhang zwischen Inflation und Output in Deutschland unter Bercksichtigung der Inflationswerwartungen," Deutsche Bundesbank Discussion Paper No. 5/97.

ROBERTS, J. M. (1997): "Is inflation sticky?," *Journal of Monetary Economics*, 39, 173–196.

RUDD, J., AND K. WHELAN (2005): "New tests of the new–keynesian Phillips curve," *Journal of Monetary Economics*, 52(6), 1167–1181.

Rumler, F. (2006): “Estimates of the open economy New Keynesian Phillips curve for Euro area countries,” *Open Economies Review*, 18(4), 427–451.

Sbordone, A. M. (2005): “Do expected future marginal costs drive inflation dynamics?,” *Journal of Monetary Economics*, 52(6), 1183–1197.

Seitz, H. (1988): “The estimation of inflation forecasts from business survey data,” *Applied Economics*, 20(4), 427–438.

Smets, F. (2003): “Maintaining price stability: How long is the medium term,” *Journal of Monetary Economics*, 50(6), 1293–1309.

Stock, J. H., J. H. Wright, and M. Yogo (2002): “A survey of weak instruments and weak identification in generalized method of moments,” *Journal of Business & Economic Statistics*, 20(4), 518–529.

Stock, J. H., and M. Yogo (2005): “Testing for weak instruments in linear IV regression,” in *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, ed. by D. Andrews, and J. H. Stock, pp. 80–108. Cambridge University Press.

Svensson, L. E. O. (1997): “Inflation forecast targeting: Implementing and monitoring inflation targets,” *European Economic Review*, 41, 1111–1146.

Taylor, J. B. (1980): “Aggregate dynamics and staggered contracts,” *The Journal of Political Economy*, 88(1), 1–23.

Theil, H. (1952): “On the time shape of economic microvariables and the Munich business test,” *Revue de l’Institut International de Statistique*, 20, 105–120.

Walsh, C. E. (2003): *Monetary theory and policy.* The MIT Press, Cambridge.

Weber, A. (2007): “Heterogeneous expectations, learning and European inflation dynamics,” Deutsche Bundesbank Discussion Paper No. 16/07.

Weber, E. (1834): *De pulsu, resorptione, auditu et tactu. Annotationes anatomicae et physiologicae.* C.F. Köhler, Leipzig.

Zhang, C., D. R. Osborn, and D. H. Kim (2006): “Observed inflation forecasts and the New Keynesian Phillips curve,” *The School of Economics Discussion Paper Series*, 79, The University of Manchester.

Curriculum Vitae

5. Dezember 1977	Geboren in Würzburg.
1997	Abitur am Schiller–Gymnasium, Hameln.
1997 - 1998	Wehrdienst in Bremen.
1998 - 2003	Studium der BWL und VWL an der Julius–Maximilians Universität, Würzburg.
2003	Abschluß: Diplom–Volkswirt an der Julius–Maximilians Universität, Würzburg.
2003 - 2008	Doktorand der Ludwig–Maximilians–Universität, München sowie wissenschaftlicher Mitarbeiter am ifo Institut für Wirtschaftsforschung an der Universität München.

München, den 25. Februar 2008